The not-so-sterile womb: New data to challenge an old dogma

Lisa Faye Stinson, B.Sc., M.MedSc.

This thesis is presented for the degree of Doctor of Philosophy of the University of Western Australia

Faculty of Health and Medical Sciences, Division of Obstetrics & Gynaecology

2019
Dedication

I dedicate this thesis to my mum, Judy, the first in her family to go to university. My incredible mother, who, with two degrees, four children, and an amazing career in her wake, showed me that I could do anything. You have always been my hero.

This thesis is also dedicated to my dad, Cliff. We all know that you’re the real scientist in the family. Thank you for giving me a fraction of your curiosity and intellect. Thank you for always asking me “why” and “how”.

But most of all, this thesis is for Sam. You have stuck by my side during these four incredibly difficult years. You have listened to me tell you “I don’t think I can do it” and you have not believed me. You’ve been the Samwise to my Frodo. Thank you for always seeing the light inside me, even when I couldn’t find it. I truly couldn’t have finished this without you.
Thesis declaration

I, Lisa Stinson, certify that:

This thesis has been substantially accomplished during enrolment in this degree.

This thesis does not contain material which has been submitted for the award of any other degree or diploma in my name, in any university or other tertiary institution.

In the future, no part of this thesis will be used in a submission in my name, for any other degree or diploma in any university or other tertiary institution without the prior approval of The University of Western Australia and where applicable, any partner institution responsible for the joint-award of this degree.

This thesis does not contain any material previously published or written by another person, except where due reference has been made in the text and, where relevant, in the Authorship Declaration that follows.

This thesis does not violate or infringe any copyright, trademark, patent, or other rights whatsoever of any person.

The research involving human data reported in this thesis was assessed and approved by The University of Western Australia Human Research Ethics Committee. Approval #: RA/4/1/8280. Written patient consent has been received and archived for the research involving patient data reported in this thesis.

The following approvals were obtained prior to commencing the relevant work described in this thesis: Human Research Ethics Committee of the Western Australian Department of Health’s Women and Newborns Health Service (2015212EW).

This thesis contains published work and work prepared for publication, some of which has been co-authored.

Signature: [Redacted]

Date: 16/01/2019
Authorship declaration: co-authored publications

This thesis contains work that has been published and prepared for publication.

Details of the work:


Location in thesis:

Chapter 2, page 12.

Student contribution to work:

LS conceived of, researched, and wrote the manuscript. Her co-authors critically edited the manuscript.

Co-author signatures and dates:

15/01/2019 16/01/2019

Details of the work:


Location in thesis:

Chapter 3, page 37.

Student contribution to work:

LS performed all experiments and data analysis, and wrote the manuscript. Her co-authors critically edited the manuscript. All authors contributed to study design.

Co-author signatures and dates:

15/01/2019 16/01/2019
Details of the work:


Location in thesis:

Chapter 4, page 63.

Student contribution to work:

LS performed all experiments and data analysis and wrote the manuscript. Her co-authors critically edited the manuscript. MP conceived of the study.

Co-author signatures and dates:

15/01/2019 16/01/2019

Details of the work:


Location in thesis:

Chapter 5, page 75.

Student contribution to work:

LS recruited all patients involved in the study and collected all samples. MB developed the SCFA analysis, and LS & MB performed this experiment. LS performed all other experiments and all data analysis and wrote the manuscript. Her co-authors critically edited the manuscript. All authors contributed to study design.

Co-author signatures and dates:

09/01/2019 15/01/2019 16/01/2019
Details of the work:


Location in thesis:

Chapter 7, page 120.

Student contribution to work:

LS performed all experiments and data analysis and wrote the manuscript. Her co-authors critically edited the manuscript. MP conceived of the study.

Co-author signatures and dates:

15/01/2019 16/01/2019

Details of the work:


Location in thesis:

Chapter 8, page 133.

Student contribution to work:

LS recruited all patients involved in the study and collected all samples. She performed all experiments and data analysis and wrote the manuscript. Her co-authors critically edited the manuscript. JK conceived of the study.

Co-author signatures and dates:

15/01/2019 16/01/2019
I, Jeffrey Keelan, certify that the student’s statements regarding their contribution to each of the works listed above are correct.

Coordinating supervisor signature:

Date: 16/01/2019
Abstract

It has long been assumed that the human fetus develops in a sterile environment. Over the last decade new data have emerged to challenge this dogma. However, these ideas are still controversial and there is disagreement amongst researchers as to whether the human microbiome is or is not seeded prior to birth. While numerous studies have now reported the detection of bacterial, archaeal, fungal, and viral DNA in meconium and amniotic fluid, much of the data that have been produced are contentious due to underlying contamination issues. Further, the biological significance of prenatal exposure to non-pathogenic microbes is largely unknown.

Here, we have collected first-pass meconium and amniotic fluid samples from 50 women delivering by elective caesarean section for bacterial profiling. The full length 16S rRNA gene was sequenced using PacBio SMRT cell technology, to allow high resolution profiling of the fetal gut and amniotic fluid bacterial microbiomes. Additionally, levels of inflammatory cytokines were measured in amniotic fluid, and placental inflammation was assessed by histology. In order to relate fetal microbial exposure to physiological outcomes, levels of immunomodulatory short chain fatty acids (SCFAs) were quantified in meconium, and fetal immune responses were measured in umbilical cord blood. Further, 1198 mid-gestation amniocentesis samples (14-20 weeks gestation) from a Swedish cohort of uncomplicated pregnancies were screened by qPCR for bacterial DNA; 18 of these were selected for 16S rRNA gene sequencing.

Meconium is a difficult substance to extract DNA from, contains high levels of PCR inhibitors, and, in normal circumstances, contains very low amounts of bacteria. We initially optimised our DNA extraction methods for this sample type to increase DNA yield and decrease the co-extraction of PCR inhibitors. We then determined that the major source of contamination in bacterial microbiome studies is the PCR master mix reagents (the “mixome”), and successfully used dsDNase treatment to decontaminate these, minimising reagent-based contamination and allowing accurate microbial profiling of low microbial biomass samples.

Using these optimised methods, we show convincing evidence that the human bacterial gut microbiome is established prior to birth. All meconium samples contained detectable levels of bacterial DNA and the immunomodulatory SCFAs acetate and
propionate, confirming the hypothesis that the fetal gut is inoculated with bacteria/bacterial DNA in-utero. Maternal type II diabetes was associated with a significantly lower level of propionate in meconium samples compared to both healthy mothers \((P = 0.005)\) and mothers with gestational diabetes \((P = 0.003)\). At the phylum level, meconium was dominated by reads that mapped to Proteobacteria. *Pelomonas puraquae* was the most abundant read identified. This genus has often been identified as a contaminant in microbiome studies; however, it was absent in 70% of our negative extraction controls, and all PCR negative controls, so an external contamination source is unlikely. Importantly, this genus has previously been identified in the core non-pregnant endometrial microbiome, but not paired vaginal samples, suggesting it was not acquired through vertical transfer. Reads associated with *Lactobacillus* sp. (which dominates the vaginal microbiome) were found in only two samples. Importantly, using propidium monoazide to differentiate between intact and non-viable cells, we determined that the majority of DNA that was recovered from these meconium samples belonged to viable bacterial cells, suggesting that the fetal gut is indeed seeded with live bacteria prior to birth. This is an important finding, as exposure to live bacteria during gestation would be expected to have a significant impact on the developing fetus via the production of microbial metabolites and activation of innate immune sensing pathways.

Most full-gestation amniotic fluid samples (83.7%) contained detectable levels of bacterial DNA. The full-gestation amniotic fluid bacterial microbiome was distinct from that of meconium, suggesting early niche differentiation, and was dominated by skin commensals. Of the mid-gestation amniocentesis samples that were screened, 238 (19.9%) were positive for bacterial DNA, significantly less than the late pregnancy samples \((P < 0.001)\); however, bacterial alpha diversity was greater in mid-gestation samples than full-gestation samples (Shannon diversity \(P = 0.006)\).

Together, these data provide new evidence that seeding of the human bacterial microbiome commences prenatally, possibly very early in pregnancy in some cases. Maternal health may influence fetal immune programming via modulation of the fetal microbiome and immunomodulatory SCFAs. Microbial niche differentiation likely begins prior to birth.
List of figures

FIGURE 1.1: Relative abundance of bacterial, fungal and viral communities at different body sites exposed to the external environment — the nose, mouth, skin, stomach, intestinal tract, vagina and lungs. Sourced from [27].

FIGURE 1.2: Stepwise bacterial colonisation of the human gut in early life. Sourced from [91].

FIGURE 2.1: The major bacterial genera reported in maternal and fetal microbial sites, and possible origins relating to fetal microbiome seeding. *Although microbiome data are not available, there are data on targeted microbiology of this environment. *S. aureus and *Candida* spp. have previously been detected in 73% and 5% of newborn oronasal swabs, respectively [111].

FIGURE 2.2: A potential mechanism for bacterial translocation from the maternal gut to the fetus via the action of dendritic cells. It remains unclear how the bacteria may spread from the placenta to the fetus, though spread through the fetal membranes and into the amniotic fluid is one possibility.

FIGURE 2.3: The points at which bacterially-derived SCFAs can interact with the immune system. Adapted from [94].

FIGURE 3.1: DNA quantitation (ng/g meconium) for each meconium sample using various extraction methods.

FIGURE 3.2: 16S rRNA gene Ct values from meconium samples (n = 5 for kits QS, QM, PS, and PM, n = 4 for kit PM2) and negative extraction controls (NEG) (n = 1 per kit) processed with various extraction kits. Sample numbers (1-5 or NEG) are displayed on the x axis. Ct values are displayed on the y axis.

FIGURE 3.3: Percent recovery of GBS DNA with meconium extractions generated from various extraction kits. Data are mean ± SD.

FIGURE 3.4: Relative abundance of OTUs in meconium samples (n=5) after extraction with various kits at phylum level.

FIGURE 3.5: Relative abundance of OTUs in meconium samples (n=5) after extraction with various kits at genus level.
FIGURE 3.6: Comparison of relative abundance of the 10 most abundant OTUs in meconium samples with extraction controls and PCR controls from A) kit QS, B) kit QM, C) kit PS, and D) kit PM. Note that abundances are relative and not absolute, so a direct comparison of abundance cannot be made between meconium samples, PCR controls and extraction controls.

FIGURE 5.1: Abundance (number of reads) of bacterial species detected in meconium samples.

FIGURE 5.2: Abundance (number of reads) of bacterial species detected in amniotic fluid samples.

FIGURE 5.3: Short chain fatty acid levels (mmol/g) in first-pass meconium samples (n = 47).

FIGURE 5.4: A: Amniotic fluid cytokine levels (pg/mL) from pregnancies with *P. acnes* DNA present (pink) or absent (blue) in the amniotic fluid. B: Amniotic fluid cytokine levels (pg/mL) from pregnancies with *P. puraquae* DNA present (pink) or absent (blue) in the amniotic fluid. *P = 0.058, **P = 0.025, ***P = 0.018.

FIGURE 5.5: A: Venn diagram showing species that were unique to meconium or amniotic fluid samples, or shared between them. B: Venn diagram showing number of reads that were unique to meconium or amniotic fluid samples, or shared between them.

FIGURE 5.6: Relative abundance of bacterial genera recovered from meconium and amniotic fluid (AF) samples.

FIGURE 5.7: Differential abundance of *P. puraquae* and *P. acnes* in amniotic fluid (pink bars) and meconium (blue bars). Absolute abundance is displayed in the left panels, and log-transformed abundance is displayed in the right panels.

SUPPLEMENTARY FIGURE 5.1: Abundance (number of reads) of bacterial species detected in meconium samples presented with the two dominant *P. puraquae* OTUs removed.

SUPPLEMENTARY FIGURE 5.2: Abundance (number of reads) of bacterial species detected in amniotic fluid samples presented with the two dominant *P. puraquae* OTUs removed.
FIGURE 6.1: Absolute abundance (number of reads) of OTUs in 18 selected PCR-positive mid-gestation amniotic fluid samples and the genera they map to (n = 18).

FIGURE 6.2: Percent abundance (highest to lowest) of bacterial species in all mid-gestation amniotic fluid samples.

FIGURE 7.1: Principle Coordinates Analysis based on weighted UniFrac distance. Individual meconium samples (n = 5) are labelled and colour coded based on whether they were PMA-treated prior to amplification and sequencing (pink) or untreated (blue).

FIGURE 7.2: Alpha diversity (observed OTUs) of meconium samples (n = 5) that were PMA treated prior to amplification and sequencing (pink) or untreated (blue).

FIGURE 7.3: The relative abundance (circle size) of the genera (y-axis) in each sample (x-axis) is plotted for PMA-treated (PMA) and untreated (UT) meconium samples.

FIGURE 8.1: Amniotic fluid cytokine levels of patients with deciduitis (DEC, n = 3), vasculitis of unknown aetiology (VUE, n = 6), chorioamnionitis (CAM, n = 9), inflammatory reaction of the fetal vessels (FIR, n = 11), or no inflammation (NI, n = 20). Data are mean ± SEM. * P < 0.05 compared to the NI group.

FIGURE 8.2: IL-8 levels in fetal cord blood plasma relative to unstimulated controls. Data is stratified according to type of placental inflammation (DEC n = 2, VUE n = 7, CAM n = 9, FIR n = 12, NI n = 23). * P < 0.05.

FIGURE 8.3: IL-10 levels in fetal cord blood plasma relative to unstimulated controls. Data is stratified according to type of placental inflammation (DEC n = 2, VUE n = 7, CAM n = 7, FIR n = 7, NI n = 19). * P < 0.05.

FIGURE 8.4: IL-1β levels in fetal cord blood plasma relative to unstimulated controls. Data is stratified according to type of placental inflammation (DEC n = 2, VUE n = 7, CAM n = 9, FIR n = 12, NI n = 23). * P < 0.05.

FIGURE 8.5: G-CSF levels in fetal cord blood plasma relative to unstimulated controls. Data is stratified according to type of placental inflammation (DEC n = 2, VUE n = 7, CAM n = 9, FIR n = 12, NI n = 23). * P < 0.05, ** P < 0.01.
FIGURE 8.6: IFN-γ levels in fetal cord blood plasma relative to unstimulated controls. Data is stratified according to type of placental inflammation (DEC n = 1, VUE n = 7, CAM n = 9, FIR n = 11, NI n = 20). * $P < 0.05$.

FIGURE 8.7: IL-6 levels in fetal cord blood plasma relative to unstimulated controls. Data is stratified according to type of placental inflammation (DEC n = 2, VUE n = 7, CAM n = 9, FIR n = 12, NI n = 23). * $P < 0.05$, ** $P < 0.01$. 
**List of tables**

TABLE 2.1: Bacteria detected in meconium samples as reported from eleven available studies. Superscripts denote genera commonly found in the gut $G$ and vagina $V$, and/or detected in the amniotic fluid $AF$ of pregnancies with intact membranes.

TABLE 3.1: Summary of DNA extraction kit characteristics.

TABLE 3.2: Summary of OTUs detected in negative extraction controls ($n = 1$ per kit) and negative PCR controls ($n = 2$).

TABLE 3.3: Summary of OTUs detected in meconium samples only, not in negative extraction controls or negative PCR controls.

TABLE 4.1: 16S rRNA gene qPCR Ct values from blank extraction controls (EC1-4) and PCR controls (NT1-2) amplified with dsDNase-treated or untreated master mix.

TABLE 4.2: 16S rRNA gene sequence reads produced from negative extraction controls (EC1-4) and negative PCR controls (NT1-2) that were amplified with or without dsDNase-treated master mix.

TABLE 5.1: Indication for caesarean delivery in this cohort ($n = 50$). Multiple indications for caesarean delivery were common.

TABLE 5.2: Maternal, fetal, and pregnancy characteristics for this cohort ($n = 50$). Values are reported as mean (range) or $n$ (percent).

TABLE 5.3: OTUs detected in blank controls taken during DNA extraction from meconium samples (Meconium extractions EC 1-5) and amniotic fluid samples (AF extractions EC 1-5). Percent homology to sequences reported on the RDP database is included. All blank PCR controls were completely negative and are therefore not included in this table.


TABLE 6.1: Pregnancy demographic characteristics of the cohort ($n = 1198$). Values are reported as mean (range) or $n$ (percent).
TABLE 7.1: Quantity of total DNA in each meconium sample with (+) or without (-) PMA treatment.

TABLE 8.1: Indication for caesarean delivery in this cohort. In many cases there were multiple indications for caesarean delivery.

TABLE 8.2: Maternal and fetal characteristics. Data are presented as median (IQR [Range]) or N (%). P-values represent analysis of inflammation present (+) vs inflammation absent (-). Abbreviations: N, number; n.s., not significant (P > 0.05); GA, gestational age; wk, weeks; IQR, interquartile range.

TABLE 8.3: Details of immune stimulants used in this study.


TABLE 8.5: Baseline levels of cytokines (pg/mL) in cord blood of patients with deciduitis (DEC, n = 2), vasculitis of unknown aetiology (VUE, n = 6), chorioamnionitis (CAM, n = 9), inflammatory reaction of the fetal vessels (FIR, n = 11), or no inflammation (NI, n = 20). Data are median ± IQR.

TABLE 8.6: Mean fold change in cord blood cytokine levels following 24 h incubation with various immune stimuli.

TABLE 8.7: Mean cord blood cytokine levels after stimulation with LPS or PGN in neonates whose meconium contained Gram negative bacteria only or a combination of Gram negative and positive bacteria, and in neonates whose amniotic fluid contained Gram negative bacteria only, Gram positive bacteria only, or a combination of Gram negative and positive bacteria.
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>5’ppp-dsRNA</td>
<td>5’ triphosphate double-stranded RNA</td>
</tr>
<tr>
<td>AF</td>
<td>Amniotic fluid</td>
</tr>
<tr>
<td>BIOM</td>
<td>Biological observation matrix</td>
</tr>
<tr>
<td>BLAST</td>
<td>Basic local alignment search tool</td>
</tr>
<tr>
<td>BMI</td>
<td>Body mass index</td>
</tr>
<tr>
<td>BV</td>
<td>Bacterial vaginosis</td>
</tr>
<tr>
<td>CAM</td>
<td>Chorioamnionitis</td>
</tr>
<tr>
<td>CCS</td>
<td>Circular consensus sequence</td>
</tr>
<tr>
<td>CD8</td>
<td>Cluster of differentiation 8</td>
</tr>
<tr>
<td>cGAMP</td>
<td>Cyclic guanosine monophosphate–adenosine monophosphate</td>
</tr>
<tr>
<td>CHILD</td>
<td>Canadian Healthy Infant Longitudinal Development</td>
</tr>
<tr>
<td>CSIRO</td>
<td>Commonwealth Scientific and Industrial Research Organisation</td>
</tr>
<tr>
<td>CXCL10</td>
<td>C-X-C motif chemokine 10</td>
</tr>
<tr>
<td>DC</td>
<td>Dendritic cell</td>
</tr>
<tr>
<td>DEC</td>
<td>Deciduitis</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>dsDNA</td>
<td>Double stranded DNA</td>
</tr>
<tr>
<td>dsDNase</td>
<td>Double stranded DNase</td>
</tr>
<tr>
<td>dsRNA</td>
<td>Double stranded RNA</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EC</td>
<td>Extraction control</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>FIR</td>
<td>Fetal inflammatory reaction</td>
</tr>
<tr>
<td>FISH</td>
<td>Fluorescent in-situ hybridisation</td>
</tr>
<tr>
<td>G-CSF</td>
<td>Granulocyte-colony stimulating factor</td>
</tr>
<tr>
<td>GA</td>
<td>Gestational age</td>
</tr>
<tr>
<td>GBS</td>
<td>Group B streptococcus</td>
</tr>
<tr>
<td>GC-MS</td>
<td>Gas chromatography – mass spectrometry</td>
</tr>
<tr>
<td>GDM</td>
<td>Gestational diabetes mellitus</td>
</tr>
<tr>
<td>GHAP</td>
<td>Greenfield hybrid amplicon pipeline</td>
</tr>
<tr>
<td>h</td>
<td>Hour</td>
</tr>
<tr>
<td>HITChip</td>
<td>The Human Intestinal Tract Chip</td>
</tr>
<tr>
<td>HIV</td>
<td>Human immunodeficiency virus</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>HS</td>
<td>High sensitivity</td>
</tr>
<tr>
<td>IAP</td>
<td>Intrapartum antibiotic prophylaxis</td>
</tr>
<tr>
<td>IBD</td>
<td>Inflammatory bowel disease</td>
</tr>
<tr>
<td>IFN</td>
<td>Interferon</td>
</tr>
<tr>
<td>IHMS</td>
<td>International human microbiome standards</td>
</tr>
<tr>
<td>IL</td>
<td>Interleukin</td>
</tr>
<tr>
<td>IQR</td>
<td>Interquartile range</td>
</tr>
<tr>
<td>IRF3</td>
<td>Interferon regulatory transcription factor 3</td>
</tr>
<tr>
<td>Acronym</td>
<td>Description</td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
</tr>
<tr>
<td>IRT</td>
<td>Inhibitor removal treatment</td>
</tr>
<tr>
<td>KF</td>
<td>King Fisher</td>
</tr>
<tr>
<td>LED</td>
<td>Light emitting diode</td>
</tr>
<tr>
<td>LGA</td>
<td>Large for gestational age</td>
</tr>
<tr>
<td>LGG</td>
<td>Lactobacillus rhamnosus GG</td>
</tr>
<tr>
<td>LPS</td>
<td>Lipopolysaccharide</td>
</tr>
<tr>
<td>MALT</td>
<td>Mucosal-associated lymphoid tissue</td>
</tr>
<tr>
<td>MEC</td>
<td>Meconium</td>
</tr>
<tr>
<td>Myd88</td>
<td>Myeloid differentiation primary response 88</td>
</tr>
<tr>
<td>ND</td>
<td>Not detected</td>
</tr>
<tr>
<td>NEC</td>
<td>Necrotising enterocolitis</td>
</tr>
<tr>
<td>NEG</td>
<td>Negative</td>
</tr>
<tr>
<td>NF-KB</td>
<td>Nuclear factor kappa-light-chain-enhancer of activated B cells</td>
</tr>
<tr>
<td>NGS</td>
<td>Next generation sequencing</td>
</tr>
<tr>
<td>NH4</td>
<td>Ammonium</td>
</tr>
<tr>
<td>NI</td>
<td>No inflammation</td>
</tr>
<tr>
<td>NK</td>
<td>Natural killer</td>
</tr>
<tr>
<td>NOD1</td>
<td>Nucleotide-binding oligomerisation domain-containing protein 1</td>
</tr>
<tr>
<td>NT</td>
<td>No template</td>
</tr>
<tr>
<td>OTU</td>
<td>Operational taxonomic unit</td>
</tr>
<tr>
<td>PacBio</td>
<td>Pacific Biosciences</td>
</tr>
<tr>
<td>PBMC</td>
<td>Peripheral blood mononuclear cell</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate buffered saline</td>
</tr>
<tr>
<td>PCOA</td>
<td>Principle coordinates analysis</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PERMANOVA</td>
<td>Permutational multivariate analysis of variance</td>
</tr>
<tr>
<td>PGN</td>
<td>Peptidoglycan</td>
</tr>
<tr>
<td>PM</td>
<td>MoBio MagAttract PowerMicrobiome DNA/RNA kit</td>
</tr>
<tr>
<td>PMA</td>
<td>Propidium monoazide</td>
</tr>
<tr>
<td>Poly (I:C)</td>
<td>Polyinosinic:polycytidylic acid</td>
</tr>
<tr>
<td>PPROM</td>
<td>Preterm premature rupture of the membranes</td>
</tr>
<tr>
<td>PS</td>
<td>MoBio Power Soil kit</td>
</tr>
<tr>
<td>PTB</td>
<td>Preterm birth</td>
</tr>
<tr>
<td>QIIME</td>
<td>Quantitative Insights Into Microbial Ecology</td>
</tr>
<tr>
<td>QM</td>
<td>QIAamp DNA Microbiome kit</td>
</tr>
<tr>
<td>QS</td>
<td>Qiagen QIAamp DNA Stool Kit Mini</td>
</tr>
<tr>
<td>R848</td>
<td>Resiquimod</td>
</tr>
<tr>
<td>RDP</td>
<td>Ribosomal Database Project</td>
</tr>
<tr>
<td>RIG-I</td>
<td>Retinoic acid-inducible gene I</td>
</tr>
<tr>
<td>RNA</td>
<td>Ribonucleic acid</td>
</tr>
<tr>
<td>RPM</td>
<td>Revolutions per minutes</td>
</tr>
<tr>
<td>SCFA</td>
<td>Short chain fatty acid</td>
</tr>
<tr>
<td>SD</td>
<td>Standard deviation</td>
</tr>
<tr>
<td>SEM</td>
<td>Standard error of the mean</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Full Form</td>
</tr>
<tr>
<td>--------------</td>
<td>-----------</td>
</tr>
<tr>
<td>SGA</td>
<td>Small for gestational age</td>
</tr>
<tr>
<td>sPTB</td>
<td>Spontaneous preterm birth</td>
</tr>
<tr>
<td>STING</td>
<td>Stimulator of interferon genes</td>
</tr>
<tr>
<td>T2DM</td>
<td>Type II diabetes mellitus</td>
</tr>
<tr>
<td>TE</td>
<td>Tris EDTA</td>
</tr>
<tr>
<td>Th</td>
<td>T-helper cell</td>
</tr>
<tr>
<td>TLR</td>
<td>Toll like receptor</td>
</tr>
<tr>
<td>TNF</td>
<td>Tumour necrosis factor</td>
</tr>
<tr>
<td>Treg</td>
<td>T regulatory cell</td>
</tr>
<tr>
<td>TRIF</td>
<td>TIR-domain-containing adapter-inducing interferon beta</td>
</tr>
<tr>
<td>UK</td>
<td>United Kingdom</td>
</tr>
<tr>
<td>US</td>
<td>Unstimulated</td>
</tr>
<tr>
<td>US-L</td>
<td>Unstimulated LyoVec control</td>
</tr>
<tr>
<td>US-R</td>
<td>Unstimulated RPMI control</td>
</tr>
<tr>
<td>USA</td>
<td>United States of America</td>
</tr>
<tr>
<td>USD</td>
<td>US dollars</td>
</tr>
<tr>
<td>UT</td>
<td>Untreated</td>
</tr>
<tr>
<td>UV</td>
<td>Ultraviolet</td>
</tr>
<tr>
<td>VUE</td>
<td>Villitus of unknown aetiology</td>
</tr>
</tbody>
</table>
Publications arising from this thesis

Journal articles arising from this thesis


Other publications arising during PhD study


Conference presentations relating to this thesis


Developmental Origins of Health and Disease Australia and New Zealand Conference. Sydney, Australia.


Acknowledgements

I would like to acknowledge my supervisors, mentors, and collaborators for their guidance during my PhD. In particular I would like to thank Professor Jeffrey Keelan for his endless support and patience. Thank you for challenging me, for providing me with so many training and travel opportunities, and for instilling in me the correct use of the word “however”. I must also acknowledge the generosity of my collaborators Professor Bo Jacobsson, Malin Barman, and Maria Hallingström for providing me with hundreds of precious amniocentesis samples. An enormous amount of work went into collecting these and I’m thrilled to be able to use them in my research. It was a complete delight to work with Dr Mary Boyce, who collaborated on this project. Mary developed a gas-chromatography mass spectrometry protocol for my meconium samples with fantastic optimism (despite my warnings that meconium was hell to work with). But above all, I would like to thank Dr Matthew Payne for being there every step of the way (and through every tumble) in my steep microbiology learning curve. Your expertise and attitude has made me into a better scientist.

My time with the UWA Division of Obstetrics and Gynaecology has been incredible, thanks largely to the wonderful people who work and study in this group. In particular, I would like to acknowledge our IT officer, Tony Smith, who brought my computer back from the dead on more than one occasion. I would also like to thank Dr Demelza Ireland for her encouragement, mentorship, and friendship over these years. It has been a pleasure to work alongside all the staff at the Large Animal Facility and King Edward Memorial Hospital.

I would like to thank every mother and every family who so generously donated their time and tissue to this study. Sharing in the birthing experiences of these women was the highlight of my PhD.

This research was supported by an Australian Government Research Training Program (RTP) Scholarship, with additional project funding by the National Health and Medical Research Council of Australia (NHMRC) and the Women’s and Infants’ Research Foundation (WIRF). I would also like to acknowledge the University of Western Australia, the Australian government and WIRF for their scholarship support.

And finally, I would like to thank chardonnay, without which I never would have gotten through this.
<table>
<thead>
<tr>
<th>Table of contents</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>DEDICATION</td>
<td>i</td>
</tr>
<tr>
<td>THESIS DECLARATION</td>
<td>ii</td>
</tr>
<tr>
<td>AUTHORSHIP DECLARATION</td>
<td>iii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>vi</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>xi</td>
</tr>
<tr>
<td>ABBREVIATION</td>
<td>xiv</td>
</tr>
<tr>
<td>PUBLICATIONS ARISING FROM THIS THESIS</td>
<td>xvii</td>
</tr>
<tr>
<td>ACKNOWLEDGEMENTS</td>
<td>xix</td>
</tr>
<tr>
<td>CONTENTS</td>
<td>xx</td>
</tr>
</tbody>
</table>

**CHAPTER 1**
The perinatal human microbiome 1
The human microbiome is indispensable to our health 2
The maternal microbiota support pregnancy adaptations and fetal immune development 3
The sterile womb hypothesis vs. the _in-utero_ colonisation hypothesis 6
Early life is a critical window for microbiome establishment 8
Research aims 9

**CHAPTER 2**
Planting the seed: origins, composition, and postnatal health significance of the fetal gastrointestinal microbiome 12
Abstract 13
Background 13
The importance of early-life microbiome establishment 15
Origins of the fetal microbiome 16
Could commensal microorganisms be selectively transferred to the fetus? 21
Microbiology of the fetal gut 21
Timing of colonisation 29
Fetal gut immunoprogramming 29
A potential role for pre- and probiotics 31
Consequences of delivery mode (caesarean vs. vaginal) 33
Antenatal antibiotic use and its effects on neonatal microbiota 35
Summary and directions for future research 36

CHAPTER 3
Comparison of meconium DNA extraction methods for use in microbiome studies 37
Abstract 38
Background 38
Methods 41
Results & discussion 46
Summary 56

CHAPTER 4
Identification and removal of contaminating microbial DNA from PCR reagents: impact on low biomass microbiome analyses 63
Abstract 64
Background 64
Methods 66
Results & discussion 69
Summary 73

CHAPTER 5
The not-so-sterile womb: Bacterial DNA profiles of the fetal gut and amniotic fluid in uncomplicated pregnancies 75
Abstract 76
Background 76
Methods 78
<table>
<thead>
<tr>
<th>Results &amp; discussion</th>
<th>84</th>
</tr>
</thead>
<tbody>
<tr>
<td>Summary</td>
<td>101</td>
</tr>
</tbody>
</table>

**CHAPTER 6**

Detection of bacterial DNA in mid-gestation amniotic fluid samples 104

Abstract 105

Background 105

Methods 108

Results & discussion 113

Summary 119

**CHAPTER 7**

Characterisation of the bacterial microbiome in first-pass meconium using propidium monoazide (PMA) to exclude non-viable bacterial DNA 120

Abstract 121

Background 121

Methods 123

Results & discussion 126

Summary 132

**CHAPTER 8**

Placental and intra-amniotic inflammation are associated with altered fetal immune responses at birth 133

Abstract 134

Background 134

Methods 136

Results & discussion 140

Summary 156

**CHAPTER 9**
Synthesis 157
Overview 158
Aim 1 158
Aim 2 160
Aim 3 163
Aim 4 164
Aim 5 166
Limitations 169
Implications 171
Directions for future research 173
The not-so-sterile womb: new data to challenge an old dogma 176

REFERENCE LIST 177

APPENDIX I 201
GHAP code

APPENDIX II 217
Chapter 1
The perinatal human microbiome

“Usually, human birth narratives focus on the origins of a new individual, focusing on the heroic travails of the mother or the amazing journey of fetus. I wish to discuss birth as the origin of a new community. For not only is the eukaryotic body being reproduced, but so also are the bodies of its symbiotic microbes and so is the set of relationships between these organic components…. The birth of the holobiont exemplifies principles of co-evolution, co-development, niche construction, and scaffolding. Birth is nothing less than the passage from one set of symbiotic relationships to another.”

The human microbiome is indispensable to our health

The human microbiome is the community of microorganisms living on and in the human body, the majority of which inhabit the intestines. These microorganisms include numerous and diverse bacteria, fungi, viruses, archaea, and protozoa. The bacterial portion of this population alone adds 38 trillion cells to the body of a healthy adult, matching the number of human cells with a ratio of approximately 1:1 [1]. These microorganisms are essential collaborators in human physiology, providing nutrient breakdown in the gut, contributing to metabolic function, calibrating the immune system, and defending against pathogens [2]. Importantly, a microbiome, by definition, is more than the sum of the microbes present. A microbiome is an entire ecological system, including the relationships that exist between microbes and their environment.

Historically, we have relied on microbial culture methods to define the bacteria that inhabit our bodies. However, due to the multitude of conditions required to successfully grow and identify all components of a microbiome, culture-depandan studies are not able to determine the true diversity of microbes present in a sample [3]. In the past two decades, culture-independent techniques such as 16S rRNA gene sequencing and shotgun sequencing have been developed and applied. Together, these techniques have extended our understanding of the human microbiome, and helped to elucidate the intimate relationship between microbes and humans. This has led to a foray of studies into the microbial ecology of various body sites (Figure 1.1), describing the role that the human microbiome plays in health and disease. Such studies have revealed that disruption to the microbiome (dysbiosis) may underpin a number of non-communicable diseases, such as asthma, allergies, obesity, cardiovascular disease, autoimmune diseases, and even psychiatric and neurologic disorders [4-9]. We have now come to think of humans not as individuals, but rather as human meta-organisms, or holobionts [10].

While the majority of studies have focussed on the composition and significance of the microbiome in mature individuals, clearly the establishment and growth of the microbiome must occur in early life and therefore has potentially profound implications for reproductive and life-long health. In the past decade a number of studies have emerged describing the role of our microbial collaborators in pregnancy and
fetal/neonatal health [4, 11-26]. While much has been uncovered in this time, many questions remain to be answered.

**FIGURE 1.1**: Relative abundance of bacterial, fungal and viral communities at different body sites exposed to the external environment — the nose, mouth, skin, stomach, intestinal tract, vagina and lungs. Sourced from [27].

**The maternal microbiome supports pregnancy adaptations and fetal immune development**

The maternal microbiome appears to play a role in programming the fetal immune system, though the mechanisms through which this occurs are unclear [28]. Nonetheless, a number of key studies have demonstrated that the maternal microbiota have marked effects on fetal immune and metabolic programming, which may have long lasting effects into infancy and childhood. In a cornerstone study of the intrauterine
microbiome by Rautava et al. [16], DNA from non-pathogenic bacteria were found in placentae and amniotic fluid of all 29 study subjects, and were associated with alterations in toll-like receptor (TLR) related innate immune gene expression in both the placenta and the fetal gut. Importantly, the authors demonstrated that this immune modulation could be altered by maternal probiotic use, thus demonstrating a link between maternal microbial composition and fetal immune function. In a double-blind placebo-controlled trial by the same authors, 241 mothers with allergic disease or atopic sensitisation received probiotics or a placebo for the final two months of pregnancy and the first two months of lactation [29]. Infants born from probiotic-treated mothers had a decreased risk of eczema in the first two years of life compared to those born of placebo-treated mothers. The combined work of Santacruz et al. and Collado et al. has shown that the composition of the maternal gut microbiome during pregnancy varies according to maternal BMI and gestational weight gain, and that such changes may impact on fetal health programming, possibly through altered serum biochemical parameters [30, 31]. Further, maternal antibiotic use during pregnancy has been associated with an increased risk of childhood asthma in the offspring, suggesting that the maternal microbiota may influence immune programming in the fetus [32].

Furthermore, there is epidemiological evidence to suggest that maternal microbial exposures might affect the future health and immune function of the offspring. Ege et al. examined 1063 children up to the age of two, surveying maternal farm animal exposure, cord blood TLR gene expression, and atopic dermatitis diagnosis in the offspring [33]. In this study, contact with farm animals during pregnancy dramatically reduced the chances of atopic dermatitis in the offspring in a dose-dependent fashion. In line with this study, others have found similar correlations between maternal farm animal exposure and a reduced risk of allergies and autoimmune disease in the offspring [23, 34-37]. This reduced risk is associated with changes in innate immune gene expression in cord blood leukocytes, suggesting that maternal microbial exposure during pregnancy can influence fetal immune development [33, 37]. However, these studies have not directly measured the maternal or neonatal microbiota; therefore, the link between maternal farm animal exposure, altered microbial exposure, and offspring disease protection requires further investigation. To further elucidate this point, Conrad and colleagues investigated the effect of maternal exposure to Acinetobacter lwofii, a bacterium isolated from cowsheds, on fetal immune physiology in a mouse model [38].
They found that *A. lwoffii* contact during pregnancy modulated TLR-gene expression in the placenta, and conferred protection against asthma in the offspring. Debarry and colleagues further demonstrated this link, showing that both *A. lwoffii* and a second farm shed isolated bacterium, *Lactococcus lactis*, reduced allergic reactions in mice [6].

Further evidence from mouse models supports the idea that the maternal microbiota shape offspring immune development. Gomez de Agüero *et al.* transiently colonised otherwise germ-free pregnant mice during gestation only with *E. coli* HA107 [39]. Pups born to transiently colonised dams had increased numbers of intestinal innate immune cells and altered intestinal gene transcription compared to those born to germ-free dams. Similarly, Nyangahu *et al.* found that antibiotic-mediated perturbations to the maternal microbiome altered offspring microbiota and immune development in mice [40]. Administration of vancomycin to pregnant dams for five days prior to delivery significantly increased the number of lymphocytes in the pups’ spleens. Administration of vancomycin in gestation also reshaped the offspring gut microbiota, with significant decreases in reads belonging to the Bacteroides phylum and significant increases in reads belonging to the Proteobacteria phylum at day 14 of life. Vancomycin was undetectable in maternal and offspring sera, suggesting that the changes in the maternal microbiota, rather than the direct action of the antibiotic *per se*, were responsible for the observed effect. In another mouse model, Gonzalez-Perez *et al.* found that maternal antibiotic administration during pregnancy and lactation resulted in dramatic alterations in the maternal and offspring gut microbiota, as well as reduced adaptive antiviral immune responses in the offspring [41]. Although offspring in this study were exposed to small quantities of antibiotics through breastmilk and trans-placental transfer, these quantities were sub-therapeutic. The authors therefore suggested that the observed effects were due to disruptions in the maternal microbiome.

Additionally, the maternal microbiome contributes to the necessary metabolic and physiological adaptations to pregnancy. Koren *et al.* reported that the maternal gut microbiota are dramatically remodelled across gestation, with first trimester stool samples closely resembling those of healthy non-pregnant women, while third trimester stool samples showed increased abundance of Actinobacteria and Proteobacteria, as well as reduced bacterial richness [12]. To explore the role that this shift in bacterial
composition plays in pregnancy, these authors transplanted faeces from women in the first and third trimesters of pregnancy into germ-free mice. The third trimester microbiota were able to induce pregnancy-like metabolism in germ-free mice, with an increase in weight gain and insulin resistance. Third trimester stool transplanted mice also had increased inflammatory responses. This study elegantly demonstrates a direct link between pregnancy-related remodelling of the gut microbiome and metabolic and immune adaptations that support pregnancy. However, it should be noted that the results of the study have not been replicated and the study has limitations due to the fact that the cohort used was not in optimal health.

Together, this body of research demonstrates the importance of the maternal microbiome for both maternal adaptations of pregnancy, and the development of the fetal immune system.

**The sterile womb hypothesis vs. the in-utero colonisation hypothesis**

It has long been assumed that the human fetus develops in a sterile environment [42]. Over the last decade new data have emerged to challenge this dogma. However, these ideas are still controversial and there is disagreement amongst researchers as to whether the human microbiome is seeded prior to birth, and, if it is, whether this low-level colonisation is biologically significant.

With some exceptions [43-45], historical, culture-based studies have been unable to detect bacteria in the amniotic fluid, placenta, or meconium from healthy, full-term pregnancies [46-48], leading to the dogma that the womb is sterile in normal healthy pregnancies. In their recent review of the evidence supporting the sterile womb hypothesis, Perez-Munoz et al. argue that the strongest evidence in favour of the hypothesis stems from the existence of germ-free animals [49]. Our ability to generate germ-free animals using transfer of fetuses from non-sterile mothers provides evidence against the existence of microbes in the placenta and fetus. However, it should be noted that commonly used germ-free animals, such as rodents and rabbits, differ greatly from humans in terms of their placentation and length of gestation. For instance, mice, which are the most commonly used germ free animal, have a gestation length of 18-22 days,
and many differences in placentation compared to humans [50]. Nevertheless, there have been rare instances where “germ-free humans” have been generated to protect neonates that are suspected of having severe immune deficiencies [51, 52]. Perez-Munoz et al. also discuss several anatomical barriers to colonisation within the human placenta, including the syncytiotrophoblast, the extravillous trophoblasts, and the basement membrane. Additionally, the presence of immune cells and anti-microbial peptides in the decidua presents an immunological barrier to microbial access to the fetus [49].

Culture-independent studies have repeatedly detected microbial DNA in first-pass meconium (a proxy for the fetal gut contents in-utero) [11, 13, 21, 24, 53-61], amniotic fluid [11, 14, 62, 63], cord blood [64], and placental samples [65-69]. However, these data are contentious due to underlying contamination issues. It is now well accepted that laboratory reagents used in microbiome studies harbour low levels of bacterial DNA [70-76]. While this contamination is not particularly problematic for studies of high microbial biomass samples, such as adult faeces, it becomes a critical issue when working with low biomass samples such as meconium or amniotic fluid. Regrettably, most of the studies in this field have not been properly controlled, casting doubt on their findings due to the possibility of contamination. Further, DNA-based sequencing studies are not able to differentiate between DNA from living bacteria and DNA from dead bacteria [77]. The biological significance of finding DNA from non-viable bacteria in the intra-amniotic cavity is unknown. DNA-based sequencing techniques are, therefore, limited in their ability to definitively confirm the existence of a true microbiome in the uterus or fetus.

While the sterile womb hypothesis is conceptually sound, culture-based evidence suggests that bacteria are able to access the fetus prior to birth in typical pregnancies. Studies dating back almost a century have been able to culture bacteria from 36-38% of first-pass meconium samples [43-45]. More recently, two groups have successfully cultured bacteria from 78.7% and 100% of first-pass meconium samples [58, 78]. Others have cultured bacteria from amniotic fluid and placental samples taken from healthy, full-term pregnancies [11]. Of course, bacterial culture is not immune to contamination, nor does it reveal the entire population of bacteria in a given sample.
These studies are, nevertheless, convincing evidence in favour of the in-utero colonisation hypothesis. Experimental evidence has also provided support for the possibility of prenatal transfer of microbes. Jimenez et al. orally inoculated pregnant mice with a genetically labelled strain of Enterococcus faecium [78]. They were able to detect this genetically labelled bacterium by PCR in meconium samples from pups delivered by sterile caesarean section, but not in meconium samples from control pups.

Prenatal transfer of microbes to the developing offspring is a universal trend in the animal kingdom. As reviewed by Funkhouser and Bordenstein, transmission of microbes to germ cells and embryos has been observed in animal phyla as diverse as Porifera, Mollusca, Arthropoda, and Chordata, suggesting an evolutionary advantage to the practice [79]. However, the authors of this review conceded that further work is required to clarify such processes in mammals.

**Early life is a critical window for microbiome establishment**

Regardless of whether the human microbiome commences prior to birth or immediately thereafter, the early-life period (that is, from conception to early childhood) is a critical window of time for the establishment and development of the microbiome. Numerous studies have demonstrated that the human gut microbiome is established in a stepwise manner throughout early infancy, with clearly defined milestones (Figure 1.2) [13, 80-82]. In particular, gut microbiome assembly progresses through three distinct stages: the developmental phase (age 3-14 months, wherein the gut microbiome is dynamic and unstable), the transitional phase (15-30 months), and the stable phase (31-46 months, in which taxonomy and diversity measures remain stable within individuals) [82]. Aberrations in the establishment of this ecosystem in infants and children underpin the risk for immune-mediated diseases such as asthma [4, 83-85]; allergies [86-90]; inflammatory bowel disease [91, 92]; metabolic conditions such as obesity [93]; and a variety of behavioural, cognitive, and mental health conditions [94, 95].

Within this early-life window, the prenatal period is of great significance. Mother-to-child transmission of bacteria during pregnancy has the potential to markedly influence postnatal health and early-life immune development. Additionally, ecological principles
tell us that the first colonisers of a naïve fetal habitat (primary successors) would likely influence subsequent colonisation events (secondary successions). As such, it is vital for us to understand potential prenatal colonisation events and their impact on the developing fetus.

![Diagram of bacterial colonisation stages](image)

**FIGURE 1.2:** Stepwise bacterial colonisation of the human gut in early life. Sourced from [96].

**Research aims**

There is convincing evidence that early-life microbial colonisation may be important to the development of the immune system and to later life health. However, the evidence regarding the nature and extent of prenatal seeding of the human microbiome remains inconclusive due to concerns around study design, confounding factors, contamination, and the interpretation of data. Differences in study methods have led to contradictory publications and disputed conclusions. Thus, there is a need to study the fetal microbiome using approaches and methodologies that will produce reliable, sensitive, and reproducible results to provide definitive findings that will conclusively address the issue.
This thesis will describe comprehensive studies of the human amniotic fluid and fetal gut bacterial microbiome based around samples collected from a cohort of pregnant women delivering by caesarean section in late pregnancy and their infants. Methodologies and approaches were developed and optimised to address the deficiencies of past studies. These results will be linked to intra-amniotic inflammation profiles, fetal immune response profiles, fetal gut metabolite production, and maternal health and lifestyle factors.

The principal aims of this research are:

**Aim 1:** To develop and validate improved methodologies for studying the fetal microbiome to create more robust and reliable data in this field.

**Aim 2:** To characterise the bacterial composition of meconium and amniotic fluid at term in normal pregnancies.

**Aim 3:** To examine the relationship between maternal health, fetal bacterial colonisation, placental inflammation, and fetal immune development.

**Aim 4:** To ascertain whether the putative fetal gut microbiome represents living, metabolically active bacterial cells, or dead cells and bacterial products transferred from the mother.

**Aim 5:** To determine whether amniotic fluid contains bacterial DNA in mid-pregnancy and, if so, to characterise the taxonomic profile of these samples.

Each chapter sets out to answer a different question relating to the overarching aims of the thesis. Chapter Two provides a review of the literature relating to fetal gut microbial colonisation, establishing strengths and weaknesses in our current knowledge. Chapter 3 compares extraction methods for use in studies of the meconium microbiome. Chapter 4 outlines a new method for eliminating reagent-based contamination in microbiome studies. Chapter 5 characterises the meconium microbiome and amniotic fluid microbiome and links these with maternal health parameters, fetal immune development, intra-amniotic inflammation profiles, and fetal gut metabolite production.
Chapter 6 explores the bacterial DNA profile of mid-gestation amniotic fluid to determine the extent to which bacterial DNA is present at this early time point in pregnancy. Chapter 7 assesses the viability of bacterial cells present in the fetal gut to ascertain whether the microbial DNA detected in these studies originated from living microbes. Chapter 8 details the immune response profiles of each neonate described in Chapter 5 and compares these responses to maternal factors, placental inflammation data, and bacterial DNA profiles. Chapter 9 draws all of this work together to critically evaluate the implications of the findings, and to make conclusions about the existence and significance of the fetal microbiome. This research will inform an expanded view of pregnancy and birth into the realm of the holobiont.
Chapter 2
Planting the seed: origins, composition, and postnatal health significance of the fetal gastrointestinal microbiota

This chapter provides a review of the literature as it stood on this topic when this project began in 2015. It highlights gaps and contradictions in the literature and proposes new avenues for research in this field. This chapter gives the necessary context for the forthcoming chapters and provides justification for the aims of this thesis. This chapter was originally published as a review article in 2016. Since its publication numerous new studies on this topic have emerged. These are discussed in later chapters of this thesis. The section on the consequences of delivery mode on the neonatal microbiome (page 33) has since been expanded upon in depth in our publication “A critical review of the bacterial baptism hypothesis and the impact of caesarean delivery on the infant microbiome” (Appendix II).

Abstract

It has long been assumed that establishment of the human microbiome commences with the birthing process. However, recent studies have found microbial DNA in umbilical cord blood, placentae, amniotic fluid, meconium, and fetal membranes in uncomplicated pregnancies, leading to the hypothesis that the seeding of the fetal microbiome may commence in-utero. In this review we will explore this hypothesis by drawing together data on the origins, dynamics, and timing of the fetal gut microbiome. The maternal origins of the fetal gut microbiota have not been conclusively determined, although bacterial translocation from the maternal circulation or ascension from the vagina, are both likely to be contributing pathways. Mother-to-child efflux of bacteria during pregnancy has the potential to markedly influence postnatal health, as the composition of gut microbiota determines production of important metabolites which are absorbed systemically and which modify immune function and development. By gaining a deeper understanding of the mechanisms underpinning fetal microbiome seeding, strategies may be developed to optimise fetal immune development and reduce the risk of adverse health and developmental outcomes.

Background

The human microbiome is the accumulation of microorganisms living on and in the human body, the majority of which inhabit the gut. This microbial community thrives largely in a relationship of mutualism within the host, providing nutrient metabolism in the gut, calibrating metabolic function, educating the immune system, maintaining community integrity, and defending against pathogens [97, 98]. Increasingly, evidence is emerging to suggest that the disruption of this community may underpin a number of non-communicable diseases, such as asthma, allergies, obesity and autoimmune diseases [4-6].

Traditionally, the establishment of human neonatal microbiota has been seen as commencing at birth; however, increasing evidence suggests that the seeding process actually occurs initially in-utero. Over a century ago, Tissier proposed that the fetus developed in the sterile environment of the womb, acquiring its microbiota from
maternal vaginal and faecal microbes during birth [42]. This became dogma, leading to the widely held belief that any microbes found in the uterine cavity must be pathological and hazardous to the fetus. For this reason most studies of fetal microbiology have come from the perspective of infection and pathology, studying abnormal pregnancies and the effects of infection on the neonate. More recently, following the “microbiome revolution”, awareness of the nature and dynamics of human microbial communities and the importance of commensal microbes to human health has increased. Evidence for the “in-utero colonisation” hypothesis is building, prompting a range of research questions regarding the origins, microbiology, and timing of this colonisation, as well as its significance in terms of immune development and health.

The acquisition of a healthy gut microbial community is vital for normal gut function and protection from disease [99, 100]. Thus, it is critical that we understand prenatal gut microbiome seeding to optimise both short- and long-term health. The fetal gut microbiota, if truly present, are likely to play an important role in fetal immunological development. Gut microbiota can modulate immune system development and function via the effects of bacterial-derived metabolites such as short chain fatty acids (SCFAs) [101-103]. Emerging evidence suggests that non-pathogenic gut bacteria may begin immunoprogramming in this way far earlier than previously expected [33, 38, 99, 100, 104-106]. There are a number of fundamental questions yet to be addressed in this area, such as when and from where is the fetal microbiome seeded, and what are the postnatal health implications of disruptions to this process? This review aims to explore these questions by critically reviewing literature on this topic.

While questions certainly remain as to the existence, viability, and biological significance of the fetal gut microbiota, the aim of this review is to discuss the evidence that supports the view that bacteria are able to access the fetus during healthy pregnancies.
Importance of early-life microbiome establishment

As the field of microbiome research has developed, a number of exciting and unexpected discoveries have been made. New evidence suggests that the fetus is not sterile, as was previously assumed. A diverse range of microbes have been detected in umbilical cord blood [64], amniotic fluid [16, 107, 108], the placenta [65-69], and the fetal membranes [16, 109] from apparently normal pregnancies without any indication of inflammation or pathology. However, it should be noted that the risk of external contamination in microbiome studies of low-biomass environments such as these is considerable, and, as such, data from these studies must be interpreted with caution.

Most studies describing the microbiota of meconium (the first stool sample, which may be used as a proxy for fetal gut contents) have identified a transient microbial community of low diversity and abundance. The dominant bacterial species identified in meconium typically reflect that of the adult gut microbiota [24, 78, 110].

Human gut microbiota are responsible for performing a range of tasks that are indispensable to the host, such as the production of vitamins, and the breakdown of oligo-/polysaccharides [111-113]. Consequently, the host must provide an environment that is sympathetic to this commensal community, while also defending against pathogenic species. This requires the host to develop an elaborate immune system for differentiating between harmful and beneficial bacteria. Education of this immune system is partially informed by early gut microbial exposure, both in-utero and postnatally.

The infant gut microbiome has been shown to influence postnatal immune development and might therefore have long-reaching health consequences. A difference in the gut microbiota of infants who will and won’t develop an allergy may be present before any clinical symptoms are evident [87, 114]. Early-life perturbations of the gut microbiota have also been postulated to be a cause of chronic inflammatory diseases such as Coeliac disease and irritable bowel disease (IBD) (reviewed by Rautava et al. [106] and Belkaid & Hand [115]). Furthermore, meconium samples with microbial communities dominated by lactic acid bacteria have been associated with risk of development of infant respiratory problems [24]. These studies suggest that long term health and disease is, in part, determined by the acquisition and maintenance of optimal microbial
communities *in-utero* and in early life. However, it should be noted that these are association studies that do not specifically address causality; high quality evidence linking early-life gut colonisation and long term health outcomes is currently lacking. Thus, it is important to understand the origins, composition, and timing of colonisation of fetal gut microbiota, and how these microbial communities change throughout the early stages of life.

**Origins of fetal microbiota**

If the fetus is seeded with commensal bacteria *in-utero*, where does this bacterial exposure come from? Present data suggests that the maternal gut, vagina, or oral cavity, are potential sources. Possible connections between the maternal and fetal microbiomes are presented diagrammatically in Figure 2.1. However, it is important to be mindful that there may not be a single source; that sources may vary between pregnancies; and that fetal seeding may be a dynamic, gestational age-dependent process.

**FIGURE 2.1:** The major bacterial genera reported in maternal and fetal microbial sites, and possible origins relating to fetal microbiome seeding. *Although microbiome data are not available, there are data on targeted microbiology of this environment. *S. aureus* and *Candida* spp. have previously been detected in 73% and 5% of newborn oronasal swabs, respectively [116].
Maternal gut microbiota

A number of recent studies have pointed to the maternal gut as a source of the fetal gut microbiota. Many of the bacterial genera found in meconium, such as Enterococcus and Escherichia, are typical of the adult gut [24, 78]. Furthermore, in their assessment of the placental microbiome, Aagaard and colleagues found E. coli, a common gut inhabitant, to be the most abundant species present [67]. Such observational evidence is supported by animal studies. These studies found that orally ingested microbial DNA was transferred to the fetus. The authors of these studies proposed a trans-placental route of bacterial transmission [117, 118]. This notion was expanded upon by Jimenez et al., who provided confirmatory evidence of maternal microbial transfer from the gastrointestinal tract. They orally inoculated pregnant mice with genetically-labelled Enterococcus faecium (isolated from human breast milk) and then examined amniotic fluid samples from term offspring after sterile caesarean section. E. faecium with the genetic label was cultured from the amniotic fluid of pups from inoculated mothers, but not from pups of control mice [64]. Later, the same group performed a similar study, this time sampling the meconium of the caesarean section delivered pups, again finding the genetically labelled E. faecium in samples from the pups of inoculated mothers, and not from those of control mothers [78]. These studies demonstrate that orally ingested live bacteria are able to access the fetal gut and the amniotic fluid in mice.

While there is good evidence that the fetus may be exposed to maternal gut microbes prenatally, the mechanisms by which microbes may be transported from the maternal gastrointestinal tract to the developing fetus have yet to be confirmed. Currently, the most likely hypothesis is that microbes are translocated from the gut epithelium into the bloodstream, and then delivered to the placenta. There are several studies that support this theory. Typically, the epithelial barrier prevents bacterial entry into the bloodstream; however, dendritic cells (DCs) have been shown to actively penetrate the gut epithelium and take up bacteria from the intestinal lumen. These bacteria-loaded DCs can then traffic to the mesenteric lymph node, via the intestinal lymphatics [119-122]. DCs can sequester live commensal bacteria here for several days [119]. Once attached to DCs and translocated to the lymphatic system, bacteria could spread to other locations throughout the body, as there is a circulation of lymphocytes within the mucosal-associated lymphoid tissue (MALT) system. Previously, antigen-stimulated
cells have been shown to be translocated from the gut to distal mucosal surfaces such as the genitourinary tract [123, 124]. Intestinal bacteria are also found in breast milk, and the same mechanism of translocation has been theorised [125, 126]. In further support of this theory, microbial translocation from the gut increases during pregnancy in mice. Perez et al. found that pregnant mice were 60% more likely to harbour bacteria in their mesenteric lymph node than non-pregnant mice, perhaps indicating an increase in DC trafficking to provide maternal microbial exposure to the fetus [125]. A model describing how maternal gut bacteria could be transferred to the fetus is presented in Figure 2.2.

**FIGURE 2.2:** A potential mechanism for bacterial translocation from the maternal gut to the fetus via the action of dendritic cells. It remains unclear how the bacteria may spread from the placenta to the fetus, though spread through the fetal membranes and into the amniotic fluid is one possibility.
**Vaginal microbiota**

The vaginal microbiome is a potential source of seeding of the fetal microbiome, due to its proximity to the uterus. It is well established that microbes may ascend from the vagina during pregnancy and invade the amniotic cavity [127-129]. Two models have been proposed describing the series of events leading to amniotic infection through this route, although the actual mechanics of this process has actually received very little experimental investigation. In the first model, bacteria ascend from the vagina and contact the decidua, then spread throughout the fetal membranes before passing into the amniotic fluid. In the second model bacteria first invade the amniotic fluid after penetrating a discrete section of the fetal membranes. The bacteria then proliferate in the amniotic fluid and eventually invade the amnion and then the chorion at multiple sites [128]. Either way, vaginal microbes are certainly able to access the amniotic fluid, which can result in colonisation of the fetal gut following fetal ingestion of amniotic fluid, and the fetal respiratory tract following fetal aspiration of amniotic fluid [130]. A large number of studies have identified DNA from vaginal microbes in amniotic fluid [131-133], fetal membranes [109, 134-136], and the placenta [67, 137] in both normal pregnancies and those with adverse outcomes. Specifically, ascending vaginal microbes have been associated with intrauterine inflammation and the pathogenesis of preterm birth (PTB). Bacteria of the class Mollicutes (in particular *Ureaplasma parvum* and *Ureaplasma urealyticum*) are typical vaginal colonisers, and are the most commonly isolated microbes from the amniotic fluid of preterm deliveries [133]. These bacteria have been shown to provoke a vigorous intrauterine inflammatory response, leading to PTB [138]. In addition, bacteria commonly associated with bacterial vaginosis (BV) are also frequently isolated from amniotic fluid extracted from preterm deliveries [139]. The association of BV and PTB has been well documented [140-143]. Interestingly, the fetal inflammatory response associated with genital mycoplasma colonisation of the placenta is clearly distinct from that of BV-associated organisms [137].

Non-pathogenic bacteria typical of the vagina, such as *Lactobacillus* spp., have also been detected in the intra-amniotic space in healthy, full-term pregnancies [16]. Further studies into the presence of non-pathogenic vaginal microbes in the intra-amniotic space are required to clarify their potential role in fetal gut microbiome seeding. Moreover, there is a lack of information regarding the timing of ascension of vaginal microbes into the intra-amniotic space, and the immune-modulating function of said microbes. It is,
however, clear that certain vaginal microorganisms are able to access the intra-amniotic space, colonise the placenta and fetal membranes, and can be involved in the aetiology of intrauterine and fetal inflammation and PTB.

*Oral cavity microbiota*

Despite its spatial separation from the intra-uterine cavity, the maternal oral cavity has been proposed as a source of fetal microbial colonisation. Previous studies have reported the detection of pathogenic oral species such as *Fusobacterium nucleatum* and *Porphyromonas gingivalis* in the placenta and amniotic of women with periodontal disease [129, 144, 145], suggesting that opportunistic migration of bacteria from the oral cavity to the intra-uterine cavity is possible. In fact, oral bacteria have been previously identified in the amniotic cavity of 13% of preterm deliveries [139]. This theory is supported by the observation that periodontal disease is associated with an increased risk of PTB [146-151]. However, interestingly, clinical trials to treat the condition during pregnancy have failed to reduce the rate of PTB [152-154]. It is worth noting here that a number of oral bacteria identified in the intra-amniotic space of women delivering preterm are also commonly found in the vagina (e.g. *Peptostreptococcus oralis*, *Streptococcus salivarius*, *Dialister* spp., and *Veillonella* spp.) and the gut (e.g. *Fusobacterium nucleatum* and *S. salivarius*) [139].

Until recently the passage of bacteria from the mouth to the intra-uterine space had only been studied in the context of oral disease and the aetiology of PTB. However, Aagaard *et al.* [67] have reported that the placental microbiome is more similar to the oral microbiome in terms of bacterial composition than any other body site tested (including the vagina and the gut). Unfortunately, the microbiota of the placenta used in this study were compared to those of body sites of unrelated, non-pregnant subjects enrolled in a separate study, which makes direct assessment of routes of transmission difficult. Nevertheless, the findings are intriguing. If the placental microbiome resembles the oral microbiome in healthy women without periodontal disease, this may imply that there is a migration of commensal bacteria in a setting of health, as well as migration of pathogens in the setting of periodontal disease.
Could commensal microorganisms be selectively transferred to the fetus?

If microbes are truly able to access the developing fetus during pregnancy, the question of whether this is a passive or active/selective process remains to be determined. Previously bacteria detected in placenta from preterm deliveries were shown to be associated with distinctive systemic inflammatory responses in the neonate [137]. Fichorova et al. measured 25 protein biomarkers of inflammation in dry blood spots taken ∼48 h after delivery from 527 caesarean-delivered preterm infants, 214 of which had culture-positive placentae. Microbes which have been previously described as being associated with PTB, such as Ureaplasma spp., Mycoplasma spp., Prevotella spp., Gardnerella vaginalis, and Peptostreptococcus spp., were each associated with a unique and strong systemic inflammatory response in the newborn, as were those which were commonly associated with BV. Conversely, and rather remarkably, the presence of Lactobacillus spp. was associated with suppression of inflammatory expression in preterm infants, both when detected alone, and in concert with other microbes. These data suggest that colonisation of the placenta by specific groups of organisms can influence the profile of circulating inflammatory proteins around the time of birth. The authors of this study employed culture techniques to detect and define microbial presence on the placenta - a method which has been shown to underestimate bacterial presence and diversity [133, 155]. This technique was utilised due to the presence of PCR inhibitors in the chorion. Importantly though, the fact that these organisms may be cultured from the placenta demonstrates that they are living and viable.

The results of this study raise a rather tantalising question: could protective bacteria be selectively translocated from the mother to the fetus as part of an active physiological process of fetal microbiome seeding and immuno-programming during pregnancy? No studies have directly addressed this possibility, but the data currently available suggest that it is scientifically plausible and worthy of investigation.

Microbiology of the fetal gut

Although many researchers have studied the development of the faecal microbiome throughout infancy and early childhood, the microbial composition of the first
meconium remains under-investigated. To date, eleven papers have been published on the topic [11, 24, 53-59, 78, 156]. The microbes present in the pre-feeding meconium can be regarded as a representation of the microbial communities of the fetal gut prior to delivery, and as such may reveal important information about intra-uterine colonisation. Unfortunately, the studies performed on meconium microbial communities to date have methodological limitations and rarely, if at all, collected metabolomics, immune, or inflammation data. Ten of the eleven studies failed to report measures taken to avoid contamination of the meconium samples from the perianal skin or the diaper. Only two of these studies included blank extraction or PCR controls, however, neither of these sequenced these controls. The majority of these same studies accepted meconium samples up to 48 h postnatally, at which point the neonatal gut would have had lengthy exposure to breastmilk or formula. Two of these studies even accepted meconium passed 3-14 days postnatally. This is understandable, especially in preterm studies, as preterm delivery can be associated with delayed passage of meconium [157]. However, by accepting meconium passed so late after birth, results may be confounded by presence of bacteria conferred to the infant gut by feeding. Breastmilk is known to be a rich source of commensal bacteria and to play an important role in seeding the infant gut microbiome. In the case of the latter, this is particularly evident in mouse-model work by Daft et al. who demonstrated that the nursing mother determined the faecal microbiota composition of the pup, not the birth mother [158]. An infant consuming ~800 mL of breast milk per day will ingest $8 \times 10^4 - 8 \times 10^6$ commensal bacteria while suckling, resulting in a faecal microbiome that closely resembles the milk microbiome [159].

Only one study to date used appropriate sampling techniques to measure microbial presence. Jiménez et al. sampled internal portions of meconium to avoid perianal contamination, and, importantly, only took pre-feeding meconium samples passed within the first two hours of life [78]. Jiménez et al. employed culture-dependent methods to examine the pre-feeding meconium passed by infants born from normal healthy pregnancies. Between one and five different species of bacteria were detected in each sample, with *E. faecalis*, *S. epidermis*, and *E. coli* being the predominant species. These species are common inhabitants of the adult gut; thus, their presence suggests a prenatal maternal-to-fetal transfer of gut bacteria. However, as this study relied on bacterial culture, it is highly likely that it underrepresents the true microbial diversity of
the meconium. Conversely, this approach confirms the presence of live bacteria and avoids the problem of PCR inhibitors in meconium. Evidence from the ten other published studies provide supportive evidence that meconium harbours a low abundance, low diversity microbiome. Mshvidadze et al. reported that 21 of 23 very preterm infant meconium samples had detectable bacterial DNA [59]. Ardissone et al. reported that 67% of term and preterm meconium samples contained bacterial DNA [55]. Hansen et al. were able to detect bacterial DNA in 10 out 15 meconium samples. The remaining seven studies found 100% of meconium samples to contain detectable bacterial DNA [11, 24, 53-56, 58, 156]. Together this evidence suggests that the fetal gut is colonised with bacteria prior to birth, although the data are confounded by the lack of methodological details regarding assessment and elimination of contaminating bacterial DNA.

Three studies have explored the meconium microbiome of healthy full-term infants. Collado et al. compared the bacterial composition of maternal faeces, the placenta, amniotic fluid, colostrum, meconium and infant faeces in 15 mother-infant pairs using culture and 16S rRNA gene sequencing [11]. They found that meconium had a distinct bacterial DNA profile that included many shared features with amniotic fluid, the placenta, and colostrum. These results may indicate that antenatal transfer of bacteria from mother to infant may involve amniotic fluid exposure and passage via the placental interface. Del Chie rico et al. employed Human Intestinal Tract Chip (HITChip) analysis of 16S rRNA gene amplicons to investigate the microbiota of the first meconium and subsequent stools of 31 infants. They described a “core” bacterial community that was common to all samples independent of delivery mode or feeding type. The authors suggested that these results might indicate a highly specialised community of seminal colonisers that is established prenatally [54]. In an effort to overcome the effects of PCR inhibitors (which are abundant in meconium) Hansen et al. used FISH to analyse the presence of selected bacteria in meconium from 15 full-term vaginally delivery neonates [57]. The probe set used covered Bifidobacterium, Bacteroides/Prevotella, Lactobacillaceae and Enterococcaceae, Enterobacteriaceae, Streptococcaceae, Staphylococcaceae, and Enterococcaceae. They were able to detect bacterial DNA in 10/15 of the collected samples, albeit at very low quantities. Each sample was positive for 2-5 of the selected probes, with one neonate dominated by Enterobacteriaceae and the rest by Bifidobacterium, Enterobacteriaceae,
Enterococcaceae and Bacteroides/Prevotella. Unfortunately, in this study the use of in-situ probes for selected bacterial genera may have introduced a negative bias into the analysis, leading to an underestimation of the microbial communities present in the meconium. On the other hand, the presence of PCR inhibitors in the meconium may have resulted in an underestimation of bacterial abundance or diversity in the other PCR-based studies.

Further studies have investigated the meconium microbiome and its association with maternal health and disease. Hu et al. utilised the first spontaneously evacuated meconium to explore the effect of maternal diabetes status on fetal gut colonisation using 16S rRNA gene sequencing [56]. They found the meconium microbiota of neonates born to mothers with diabetes was enriched with bacteria commonly found in the faecal microbiota of diabetic patients. This may indicate maternal-to-fetal efflux of gut bacteria, which may be affected by maternal health status. Gosalbes et al. used similar methods to evaluate the effects of lifestyle variables and atopic disease on the meconium microbiota [24]. Bacterial DNA was recovered from all meconium samples, and two clusters of bacterial communities were identified. The first was characterised by a low diversity of bacteria, but a high prevalence of enteric bacteria, and was associated with a maternal history of atopic eczema. The second was characterised by a richer diversity of bacteria, dominated by lactic acid bacteria, and was associated with neonatal respiratory illness.

Other studies have focused instead on the meconium microbiome of preterm neonates. Moles et al. compared bacterial DNA profiles of first-pass meconium with that of faecal samples collected throughout the first three weeks of life to describe the maturation of the early gut bacterial communities of preterm infants [58]. These authors used a combination of culture and molecular techniques, including HITChip analysis of 16S rRNA gene amplicons. They found a shift from Bacilli and other Firmicutes dominating meconium samples to Proteobacteria in later faecal samples. Mshvildadze et al. compared meconium bacterial DNA profiles of preterm neonates with and without necrotising enterocolitis (NEC) using 16S rRNA gene sequencing [59]. They reported that overall bacterial profiles of meconium samples from neonates who went on to develop NEC were not distinguishable from healthy controls. Lower gestational age at
delivery was associated with lower bacterial diversity. Recently, Ardissone and colleagues [55] sought to determine whether there was an association between meconium bacterial exposure and PTB. The authors investigated the meconium microbiota of 52 infants born between 23 – 41 weeks gestation, 15 of which were diagnosed with chorioamnionitis. In their study, 67% of meconium samples contained bacterial DNA. They found that meconium bacterial DNA profiles were most similar to those of the amniotic fluid, as opposed to those of the maternal vagina or oral cavity. They found a non-statistically significant trend for increased meconium bacterial DNA in early preterm infants (<33 weeks) vs. late preterm/full-term (>33 weeks), contrary to what had previously been reported by Mshvildadze et al. [59]. This is, however, consistent with the well-established inverse relationship between intra-amniotic infection and gestational age in preterm deliveries. Gestational age was found to have a significant effect on microbial community structure of meconium, with prevalence of Proteobacteria and Firmicutes being negatively correlated with gestational age. Specifically, infants delivered <33 weeks had a significantly greater proportion of 16S rRNA gene sequence reads for Enterococcus sp., Enterobacter sp., Lactobacillus sp., Photobacteriadus sp., and Tannerella sp. The authors did not explore associations between intra-uterine inflammation and meconium bacterial DNA profiles. Chernikova et al. investigated the effect of maternal complications (choorioamnionitis and prolonged preterm premature rupture of membranes (PPROM)) on the gut microbiota of nine extremely preterm infants [53]. Weekly stool samples were taken beginning with the first-pass meconium. Unsurprisingly, it was found that meconium samples from infants of pregnancies complicated by chorioamnionitis or PPROM were enriched with potentially pathogenic bacteria (such as Serratia sp. and Parabacteroides sp.). Interestingly, antibiotic treatment did not correct this imbalance, although it should be pointed out that the presence of bacterial DNA does not necessarily reflect live organisms. The gut microbiota of preterm infants has been further investigated in a small scale (n = 6) prospective study performed by Madan et al. They found that the infants who went on to develop sepsis had less bacterial diversity in their meconium than those who remained healthy.

In summary, there is consensus in the literature that the meconium microbiome has a unique constitution, and is dominated by bacteria of the Proteobacteria and Firmicutes phyla (see Table 2.1). There is also agreement amongst studies that adult gut bacteria
are frequently found in meconium samples and that meconium and amniotic fluid samples share many microbiological features. Only three of the previously described studies identified cases of sterile meconium, demonstrating that the gut is likely to be colonised \textit{in-utero}, but perhaps not in all cases. In some cases the presence of PCR inhibitors in meconium may have masked bacterial presence, especially as DNA yields from meconium are inherently low. Alternatively, contamination from the neonate’s nappy and perianal skin or from laboratory reagents may have contributed to the purported meconium “microbiome”. Studies that include meconium and later postnatal stool samples suggest that the bacterial composition of meconium influences the gut microbiome long into infancy [24, 58, 59]. However, it is difficult to make robust conclusions from these studies, due to methodological flaws and the risk of external contamination in these samples. As meconium is a low DNA yield sample, contamination from external sources, from handling, or from non-sterile reagents can greatly influence a study’s results. Due to the low-biomass nature of this sample type, blank extraction controls and no template PCR controls are critical to ensuring robust results. None of the studies presented here adequately described the measures they had taken to account for these problems, and so they must be interpreted with caution. Furthermore, a recent study has questioned the reliance of faecal microbiota as a proxy for the infant gut microbiota. Romano-Keeler and colleagues reported distinct microbial niches in infant faecal microbiota and intestinal tissue sites, suggesting that infant stool does not necessarily provide an accurate representation of the microbiota of the intestinal environment overall [160]. It should also be noted that these studies are only able to draw associations between the fetal gut microbiota and health outcomes and are not able to establish cause/effect.
TABLE 2.1: Bacteria detected in meconium samples as reported from eleven available studies. Superscripts denote genera commonly found in the gut (G) and vagina (V), and/or detected in the amniotic fluid (AF) of pregnancies with intact membranes.

<table>
<thead>
<tr>
<th>Phyla</th>
<th>Genera</th>
<th>Study identified</th>
<th>Preterm</th>
<th>Full-term</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Proteobacteria</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Acinetobacter G, V, AF</td>
<td>1, 8, 9, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Aeromonas G</td>
<td>1, 9, 10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Alcaligenes G</td>
<td>2</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Arcobacter</td>
<td>1</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Burkholderia</td>
<td>10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Campylobacter G</td>
<td>[55], 9, 10</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Comamonas</td>
<td>1, 2</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Citrobacter V, AF</td>
<td>1, 3, 4, 8, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Cupriavidus</td>
<td>2</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Deftia AF</td>
<td>2, 9, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Enhydrobacter G, AF</td>
<td>1, 9</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Enterobacter G, V, AF</td>
<td>1, 3, 4, 5, 8, 9, 10, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Escherichia G, V, AF</td>
<td>1, 3, 5, 8, 9, 10, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Haemophilus AF</td>
<td>1, 3, 9, 10, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Klebsiella V</td>
<td>1, 5, 6, 8, 9, 10, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Kluyvera G</td>
<td>8, 11</td>
<td>Yes</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Lachnobacillus G</td>
<td>10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Massilia G, AF</td>
<td>9</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Methylobacter G, AF</td>
<td>9, 10, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Moraxella</td>
<td>9, 10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Neisseria AF</td>
<td>3, 9</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Novosphingobium G, AF</td>
<td>9, 10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Ochrobactrum</td>
<td>9</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Pantoea</td>
<td>8, 11</td>
<td>Yes</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Plesiomonas G</td>
<td>11</td>
<td>Yes</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Providencia</td>
<td>11</td>
<td>Yes</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Pseudomonas G, V, AF</td>
<td>1, 9, 10, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Ralstonia AF</td>
<td>8</td>
<td>Yes</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Salmonella G</td>
<td>11</td>
<td>Yes</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Serratia</td>
<td>1, 4, 8, 10, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Shigella G, AF</td>
<td>1, 8, 9</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Sphingomonas G, AF</td>
<td>9, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Stenotrophomonas</td>
<td>2, 8, 9, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Vibrio G</td>
<td>10, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td><strong>Firmicutes</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Acidaminococcus G</td>
<td>2</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Aerococcus</td>
<td>9, 10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Anaerococcus G, AF</td>
<td>4, 9</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Anaerotruncus G</td>
<td>10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Bacillus G, AF</td>
<td>[55], 10, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Blautia G, AF</td>
<td>9, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Butyrvibrio G</td>
<td>10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Catenibacterium G</td>
<td>2</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Clostridium G, V, AF</td>
<td>1, 8, 9, 10, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Coprobacillus</td>
<td>10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Coprocococcus G</td>
<td>2</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Dialister V</td>
<td>2, 10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Dorea G</td>
<td>2, 10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td>Bacteroidetes</td>
<td>Alistipes</td>
<td>10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td>Bacteroides</td>
<td>Alloprevotella</td>
<td>9</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Bacteroides</td>
<td>2, 3, 5, 7, 8, 10</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Butyrivibrio</td>
<td>2</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Capnocytophaga</td>
<td>3, 9</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Parabacteroides</td>
<td>1, 2, 5, 8, 10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Porphyromonas</td>
<td>3, 9</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Prevotella</td>
<td>1, 2, 3, 7, 9, 10</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>Actinobaculum</td>
<td>9</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Actinomyces</td>
<td>3, 9, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Atopobium</td>
<td>1, 2, 3</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Bifidobacterium</td>
<td>2, 3, 5, 7, 10</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Brevibacterium</td>
<td>9</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Collinsella</td>
<td>2, 9, 10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Corynebacterium</td>
<td>1, 3, 4, 8, 9, 10, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Dermabacter</td>
<td>4</td>
<td>Yes</td>
<td>ND</td>
</tr>
<tr>
<td></td>
<td>Eggerthella</td>
<td>9, 10</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Gardnerella</td>
<td>3</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Gordonia</td>
<td>2</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Kocuria</td>
<td>2</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Micrococcus</td>
<td>9</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Modestobacter</td>
<td>2, 9</td>
<td>ND</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Propionibacterium</td>
<td>3, 8, 9, 10, 11</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Rothia</td>
<td>5, 9</td>
<td>ND</td>
<td>Yes</td>
</tr>
</tbody>
</table>
Timing of colonisation

The timing of fetal microbiome seeding remains an important clinical and scientific issue that has yet to be clarified. Differences in the perinatal microbiota of preterm and full-term infants have proved difficult to demonstrate, as extremely preterm infants are subject to other major microbiome disruptions such as broad spectrum antibiotic use, delayed feeding, and limited skin-to-skin contact. Koren et al. described changes in the maternal gut microbiota throughout pregnancy, and found the neonatal gut microbiota to most closely resemble that of the first trimester maternal gut [12]. This may indicate that maternal-to-fetal efflux of microbes occurs in early pregnancy, though this association by no means demonstrates a direct link. A possible source of the fetal gut bacteria is the amniotic fluid. Ureaplasma spp. have previously been detected in amniocentesis samples at 20 weeks [155, 161], Fusobacterium spp. at 21 weeks [155], and Ureaplasma parvum, Citrobacter koseri and viridans group streptococcus at 22 weeks [155, 161]. However, while pathogenic species appear to be able to access the intrauterine cavity early in gestation, this appears to be uncommon in normal pregnancy and it remains unclear when seeding of commensal bacteria in the fetus may typically occur.

Fetal gut immunoprogramming

Humans rely on their gut microbiota to ferment and breakdown insoluble dietary fibre to produce short chain fatty acids (SCFA) such as acetate, butyrate, and propionate. These metabolites influence immune function by reacting with metabolite sensing G-
protein-coupled receptors (GPRs) [101]. These receptors, found primarily on immune cells and gut epithelial cells, facilitate anti-inflammatory effects [101] (Figure 2.3). Additionally, some bacterial metabolites are responsible for regulating transcription of immune genes via epigenetic mechanisms. SCFAs can modulate the production of regulatory T cells (Tregs) by directing inhibition of promotor demethylation and/or histone deacetylases [101-103]. Commensal gut bacteria are also able to influence the immune system through their production of secondary bile acids, amino acids, vitamins, and fatty acids [162].

Importantly, microbial metabolites can have both local and systemic effects. Fibre is fermented by commensal bacteria in the distal colon to produce SCFAs, which have been shown in a number of studies to be distributed and act systemically. In one recent study, propionate was shown to affect dendritic cells and macrophages in bone marrow, as well as T helper 2 (Th2) cells in the airways [163]. In the mouse model, peptidoglycan (another bacterial product synthesised from dietary fibre, and also the major cell wall component of Gram positive bacteria) can enter the blood and prime the innate immune system, prompting the destruction of pathogenic bacteria [164]. Such studies illustrate the profound connection between a healthy gut microbiome and many types of immune responses throughout the body. Indeed, it is now becoming clear that the immunomodulatory effects of the gut microbiota are, at least in part, mediated by their production of SCFAs [102, 103, 165, 166]. This is exemplified by the fact that allergic disease is associated with altered faecal SCFA profiles [89]. Dysbiosis of the infant intestinal microbiome has been repeatedly associated with allergic disease (reviewed in [167]), though it is now becoming clear that deficiencies in Th1 and Treg function, which predispose to allergy, are already present at birth [168-171]. Thus, if fetuses are exposed to live bacteria in-utero, it is important to understand the metabolic activity of these bacteria in order to ascertain their role in fetal immunoprogramming [172].

Additionally, maternal microbial metabolites may have a direct effect on the fetal immune system. Bacterial metabolites can cross the placenta, affecting fetal immune development prior to delivery [101]. In a mouse model of pregnancy, Blumer and colleagues demonstrated that increased bacterial richness in the maternal gut led to
decreased inflammation and risk of allergy in the offspring [173]. In humans, maternal serum levels of propionate correlate with newborn length and weight [174]. These studies demonstrate the effect of maternally derived microbial metabolites on the developing immune system of the fetus, and highlight the potential benefits of a dietary fibre rich diet during pregnancy. Maternal dietary manipulation might be able to modulate the fetal microbiome, gut metabolite profiles and predisposition to systemic immune dysfunction.

**FIGURE 2.3:** The points at which bacterially-derived SCFAs can interact with the immune system. Adapted from [101].

**A potential role for pre- and probiotics**

Given that exposure to maternal microbes *in-utero* may influence the composition of the microbial community of the fetal gut, which in turn may have long lasting effects on the immune system and risk of developing certain diseases later in life, prenatal maternal oral probiotic therapy may represent a cheap, safe, and effective point of intervention for disease prevention. A number of randomised clinical trials have been conducted in recent years, demonstrating that prenatal probiotic intervention can lower the risk of allergy and atopic disease [175-179]. Rautava *et al.* found that maternal dietary probiotic supplementation significantly modulated the expression of TLR-related genes.
in the placenta and fetal gut [16]. Together with the studies showing efflux of microbes from the maternal gut to the fetus ([64, 78, 117, 118]), Rautava et al.’s work suggests that maternal probiotic therapy may be able to modulate fetal bacterial exposure, and may even offer a means to reduce the incidence of major non-communicable disease challenges, such as immuno-inflammatory disorders. However, this study did not include data on the effect of probiotics on the ecology of the fetal gut microbiota. In an informative study by Lahtinen et al. [180], administration of the probiotic strain Lactobacillus rhamnosus GG (LGG) to mothers during late pregnancy increased colonisation by Bifidobacterium spp. in the infant gut, but did not result in infant colonisation with L. rhamnosus. The mechanism by which probiotics may influence offspring seeding with other bacteria, while not colonising the gut themselves, remains to be elucidated. It may be that the maternal gut bacteria are able to influence fetal bacterial exposure through the action of bacterial metabolites, rather than through direct microbial exposure.

Maternal prenatal prebiotic treatment has also been shown to influence the microbiota of the infant gut. Prebiotics are non-digestible food items that promote the growth of beneficial microorganisms in the intestines. In a mouse model, Fujiwara et al. demonstrated that prenatal maternal supplementation with a fructo-oligosaccharide prebiotic altered the intestinal microbiome of the offspring, and also conferred protection against skin inflammation [181]. However, in a randomised, double-blind, placebo-controlled human trial, prebiotics were shown to have a bifidogenic effect on maternal gut microbiota that was not transferred to the offspring [182]. Again, this may infer that the influence of the maternal gut microbiota on the fetal gut is not only one of direct bacterial exposure, but rather of modulation of metabolites. It may also suggest that maternal prebiotic administration alone is not sufficient to influence fetal microbial seeding in humans. A combination of probiotics and prebiotics may be preferable to the use of either alone. Further research is needed to clarify the effectiveness of pre- and probiotic supplementation during pregnancy, and to understand the mechanisms by which maternally administered pre- and probiotics influence infant gut microbiota. While evidence of microbiome-mediated health benefits in pregnancy from traditional probiotics is currently conflicting, next generation probiotics may have a greater impact and may also serve to reveal more about the health implications of maternal-fetal microbial contact.
**Consequences of delivery mode (caesarean vs. vaginal)**

There is currently much emphasis, both in scientific and lay circles, on the importance of delivery mode in seeding the neonatal microbiome. It has been shown in a small study that neonates born from vaginal deliveries \((n = 4)\) have bacterial DNA profiles (gut, skin, and oral) that resemble that of their mother’s vagina (\textit{Lactobacillus} spp. and \textit{Prevotella} spp. dominated), while those born by caesarean section \((n = 6)\) have bacterial DNA profiles that resemble that of their mother’s skin and lack vaginal bacteria [183]. These authors’ data suggest that bacteria are vertically transmitted from the maternal vagina to the neonate during delivery. However, this study is confounded by the administration of antibiotics to women prior to caesarean delivery, as well as the lack of negative controls. Women undergoing caesarean sections receive intrapartum antibiotics (IAP), which may account for some of the differences observed between caesarean delivered and vaginally delivered infants. A later study by the same group trialled a perinatal intervention for restoring vaginal microbiota in caesarean delivered infants [184]. The authors exposed a subset of caesarean delivered infants \((n = 4)\) to their mother’s vaginal fluids immediately following birth and compared samples from anal, oral and skin sites to unexposed caesarean delivered infants \((n = 7)\) and vaginally delivered infants \((n = 7)\). They reported that they achieved partial restoration of the vaginal microbiota in exposed caesarean delivered infants, interpreting this to mean that vaginal microbial exposure at birth contributes to the structure of the infant microbiome. However, this study was particularly problematic methodologically, with a lack of negative controls, lack of antibiotic exposure in the vaginally delivering group, and a high level of baseline inter-individual variation. Furthermore, data presented in the supplementary material showed that only 3/7 vaginally delivered neonates had skin microbiomes dominated by \textit{Lactobacillus} spp. immediately following birth, while 1/2 caesarean delivered neonates sampled at birth also had a \textit{Lactobacillus} spp. dominated skin microbiome, as did all 3/3 of the swabbed neonates sampled at birth. The interpretation of data by these authors is therefore not scientifically sound.

A recent hallmark study performed by Backhed \textit{et al.} investigated the dynamics of the infant gut microbiome over the first year of life in a cohort of 98 mother-infant pairs [81]. Within this cohort 15.3\% of births were via caesarean section. The authors found that the faecal microbiota of these infants was enriched with operational taxonomic units (OTUs) similar to the maternal oral and skin microbiota and also to microbes
found in the surrounding environment during delivery. The faecal microbiota of infants born via vaginal delivery was dominated by microbes from the genera *Bacteroides*, *Bifidobacterium*, *Parabacteroides*, and *Escherichia/Shigella*. The early differences seen in vaginally and caesarean delivered infants gradually decreased throughout the 12 month study period; however, the gut microbiota of caesarean delivered infants remained more heterogeneous compared to the vaginally delivered infants.

Current epidemiological data indicate a higher rate of atopic disease in infants born from caesarean deliveries than their vaginally delivered counterparts [185-188]. This has led to the hypothesis that passage through the birth canal is an important step in the colonisation of the human microbiome. However, conflicting evidence in the literature has led to some contention on this point [55, 56, 59, 64, 67, 183]. Indeed, previous studies have suggested that transit through the vagina at birth has little effect on neonatal microbial colonisation [64, 189]. A study of vertical transmission of vaginal Lactobacilli from mothers to newborns showed that only approximately one quarter of infants acquire vaginal Lactobacilli from their mothers at birth, and that the acquired Lactobacilli do not last long in the intestine of the infant, as they are quickly replaced by Lactobacilli from breast milk [190]. Lactic acid-producing bacteria and other commensals have been isolated from meconium obtained from healthy neonates born by both vaginal and caesarean section-deliveries [189].

Of the eleven published papers describing the microbiota of the neonatal meconium, three have reported on differences associated with mode of delivery. Ardissone *et al.* found that mode of delivery had an effect on meconium microbial structure (R = 0.100; \( P = 0.044 \)) [55]. Conversely, both Hu *et al.* and Mshvildadze *et al.* reported no significant differences in the meconium microbiota of vaginally vs. caesarean delivered infants [56, 59]. Similarly, in their assessment of the placental microbiota, Aagaard *et al.* detected no differences in the bacterial composition of placentae delivered vaginally or by caesarean section [67]. The data on this issue are, therefore, equivocal. Clearly, further human studies are needed to clarify the impact of mode of delivery on the neonatal microbiome in both the short and long term.
Antenatal antibiotic use and its effects on neonatal microbiota

Maternal antibiotic use before, during and after birth, perturbs the composition of the offspring microbiota [191-195]. While in many cases the use of antibiotics in pregnancy is critical for health and survival, it also disrupts the natural host-microbe interactions that may contribute to perinatal microbiome seeding. The long term health consequences of this disruption have been the subject of concern and investigation. IAP is suspected to eradicate commensals and favour colonisation of the infant with antibiotic-resistant bacteria. Arboleya and colleagues recently reported that IAP had an equal or greater effect on the gut microbiota of preterm infants than direct administration of antibiotics to the infant [196]. However it should be noted that the antibiotic agent used and dosage varied between subjects. Similarly, Azad et al. reported significant changes in the gut microbiome of infants born to mothers receiving IAP. The authors studied a sub-group of 198 mother infant pairs from the CHILD study and found substantial and persistent differences in the gut microbiota of infants born from mothers exposed to IAP compared to infants born to mothers who were not exposed to IAP [197]. Specifically, IAP was associated with a decrease in Bacteroides and Parabacteroides and an increase in Enterococcus and Clostridium. Antibiotic type and dosage were not reported in this study, thus it is likely that the results of the study were complicated by variable antibiotic administration regimes.

Maternal antibiotic exposure has been associated with decreased bacterial diversity and a lower abundance of Lactobacilli and Bifidobacteria in the neonatal gut [191, 193, 198-200]. Furthermore, maternal antibiotic administration causes dysbiosis of the vaginal microbiota prior to birth, characterised by a reduction in Lactobacillus spp. in favour of an antibiotic-resistant polymicrobial environment (including Citrobacter spp., Enterobacter spp., and E. coli) [20, 198, 200-202]. These microbes might then be vertically transmitted to the offspring, affecting early colonisation events and influencing the neonatal microbiota, even after the termination of treatment. Perinatal antibiotic exposure has been associated with a number of serious morbidities, including asthma [32, 195, 203-206], necrotising enterocolitis [207, 208], and obesity [209]. These detrimental effects demonstrate the importance of minimising unnecessary antenatal and perinatal antibiotic use. Further investigation is needed to dissect variations in the effects exerted by different types of antibiotics, and different administration strategies.
Summary and directions for future research

Humans have co-evolved with their gut microbiota for millennia. Recent developments in microbial community research have highlighted our intimate relationship with these microbes. Increasing evidence now suggests that the fetal gut microbiota are seeded *in utero*, albeit transiently, subsequently affecting the microbial composition of the infant gut. Importantly, these microbes may begin priming and educating the fetal immune system far earlier than previously thought, in part via the action of metabolites such as SCFAs. Researchers have begun investigating the composition of the meconium microbiota to learn more about the events of fetal microbial seeding; however, this area of research is still in its infancy and many questions remain unanswered.

The timing of fetal gut colonisation has yet to be accurately defined and its significance determined, particularly in terms of the role of fetal immune ontogeny and programming. To date, no studies have investigated the relationship between the fetal gut microbiome and metabolome. Such studies are critical to understanding the interaction between the gut microbiota and the developing immune system. The consequences of disruptions to the fetal gut microbiota seeding process should also be explored; for example, the effects of intra-amniotic infection, preterm birth, maternal antibiotic use, maternal diet, and maternal metabolic disease.

From this novel field of research will come a better understanding of how the human meta-organism is formed, leading to development of strategies to restore or optimise early-life immune system-microbiota interactions and improve short- and long-term health benefits in at-risk infants.
Chapter 3

Comparison of meconium DNA extraction methods for use in microbiome studies

This chapter addresses Aim 1 of this thesis – to refine the present methodologies for studying the fetal microbiome. Choice of DNA extraction method is known to have an impact on the outcome of microbiome studies. Meconium is a problematic sample to efficiently extract DNA from, and is further complicated by co-extraction of PCR inhibitors. In this chapter four commonly used extraction kits are compared in order to define the optimal extraction method for use in meconium microbiome studies.

Abstract

Increasing evidence suggests that the establishment of human gut microbiota commences initially in-utero. Meconium – the first faecal material passed after birth - can be used to study fetal gut contents; however, processing meconium samples for microbiome studies presents significant technical challenges. Meconium hosts a low biomass microbiome, is tar-like in texture and contains high concentrations of PCR inhibitors. This study aimed to evaluate four different DNA extraction methods to elucidate the most effective method for bacterial DNA recovery and sequencing analysis from first-pass meconium. Samples from five infants were collected and processed using the following extraction kits: 1) Qiagen QIAamp DNA Stool Mini (QS); 2) Qiagen QIAamp DNA Microbiome (QM); 3) MoBio PowerSoil (PS); 4) MoBio MagAttract PowerMicrobiome (PM). Additionally, Kit PM was employed with a double inhibitor removal treatment step (PM2). Bacterial DNA recovery was assessed by qPCR. Levels of PCR inhibitors were measured by spiking DNA eluates with 0.1 ng of pure Streptococcus agalactiae (GBS) DNA followed by qPCR quantitation. Kit PM yielded the highest average total DNA yield (79.3 ng per gram of meconium). Samples extracted with kit PS had the highest detectable levels of 16S rRNA gene by qPCR. The ability of each kit to overcome PCR inhibition varied, with qPCR on GBS-spiked DNA from kits QS, QM, PS, and PM recovering 87.1%, 91.0%, 88.8% and 37.9% GBS DNA, respectively. Double inhibitor removal treatment improved the performance of kit PM, increasing GBS recovery to 56.5%. However, once DNA yield was normalised to the level recovered with the other kits 100% of GBS DNA was detected, suggesting that levels of PCR inhibitors are related to DNA yield from kit PM. Ion Torrent 16S rRNA gene sequencing revealed a high level of inter-kit variation in meconium microbiome structure. In particular, kit QM showed a bias towards extracting Firmicutes DNA, while the other kits extracted primarily Proteobacteria DNA. Choice of extraction kit greatly impacts on the ability to extract and detect bacterial DNA in meconium and on the microbiome community structure generated from these samples.

Background

Traditionally, the establishment of human gut microbiota has been seen as commencing at birth; however, increasing evidence suggests that the seeding process actually occurs initially in-utero [11, 210]. The gut microbiome plays a vital role in host health, with
increasing evidence emerging that the disruption of this community may underpin a number of non-communicable diseases [4-6]. Aberrations to the early-life gut microbiota may underpin the risk of asthma [4, 83], allergies [90, 211], and Crohn’s disease [212] later in life. To understand the formation of this microbial community, we must understand the origin and composition of the gut microbiome at birth, as this community may influence later colonisation patterns via the founder effect. Additionally, the fetal gut microbiota may play a role in prenatal immune programming [18, 24, 56, 156, 213].

First-pass meconium can be analysed as a non-invasive method of assessing fetal gut contents; however, there are a number of problems with processing this sample type. Meconium hosts a diverse, but low biomass microbiome [11, 24, 54, 56, 78]. As a point of comparison, previously published work suggests that meconium yields 0.2 ± 0.4 ng of prokaryotic DNA per mg of meconium, compared with 16.6 ± 6.4 ng of prokaryotic DNA per mg of stool at one year of age [13]. The low yield of bacterial DNA from meconium is further complicated by its high concentrations of PCR inhibitors [57, 214]. Meconium is a unique substance, and not stool in the traditional sense. It is not the excretion of waste products from digestion, but an accumulation of bile acids, pancreatic secretions, epithelial cells, and the residue of swallowed amniotic fluid. Meconium begins to form at the end of the first trimester of pregnancy, and is usually expelled by the infant within its first postnatal days (although in some cases the first meconium is passed before or during birth). Although the PCR inhibitors present meconium have never been isolated and identified, they are likely to include bile salts and acids (which are known to be strong inhibitors of PCR reactions [215]), glycolipids [216] (which mimic the structure of nucleic acids), and urea originating from the amniotic fluid (which degrades polymerases [217]). A previous study has shown that PCR recovery of bacterial DNA from meconium can be as low as 10% [57]. Additionally, meconium is tar-like in texture and difficult to dissolve, adding further barriers to efficient DNA extraction. Thus, it is imperative to optimise and standardise DNA extraction methods for meconium samples.

While there is an overarching agreement in the literature that the meconium microbiome has a unique constitution, dominated by bacteria of the Proteobacteria and Firmicutes
phylla, there is widespread lack of agreement in studies regarding the abundance and composition of meconium microbiota. In particular, the percent of colonised vs. sterile meconium varies greatly from study to study. Some authors have found 100% of meconium samples studied to be colonised, while others have found as little as 67% [24, 55-57, 78]. Hansen et al., in their study of the meconium microbiome in a cohort of 15 neonates, found that they were only able to recover bacterial DNA from 1 patient using PCR. To confirm the sterility of the other samples, fluorescent in situ hybridisation (FISH) was performed with probes specific for Bifidobacterium, Bacteroides-Prevotella, Lactobacillaceae / Enterococcaceae, Enterobacteriaceae, Streptococcaceae, Staphylococcaceae and Enterococcaceae. Their FISH analysis revealed that 10 of the supposedly sterile samples were in fact colonised by 2 - 5 families of bacteria. This study in particular highlights the difficulties researchers face in analysing the meconium microbiota by PCR to produce meaningful, unbiased and reproducible results.

Previous studies have compared commercially available DNA extraction kits for use in extracting bacterial DNA from stool for microbiome analysis [218-221]. However, a comparable analysis for meconium is lacking. The International Human Microbiome Standards (IHMS) consortium provides two standardised protocols for extraction of microbial DNA from stool samples, including a modified protocol for the Qiagen QIAamp DNA Stool Kit [222]. No standard operating procedures have been published for microbiome work on meconium. Given the unique qualities of this sample type, and the inherent problems with DNA yield and external contamination of low biomass samples, there is a need for the development of a meconium-specific standard protocol.

The Qiagen QIAamp DNA Stool Kit Mini (kit QS) and the MoBio Power Soil kit (kit PS) are widely used for DNA extraction from stool. More recently, both Qiagen and MoBio have released microbiome kits (QIAamp DNA Microbiome kit (kit QM) and MoBio MagAttract PowerMicrobiome DNA/RNA kit (kit PM)). Kit QM can be used to selectively recover prokaryotic DNA for microbiome analysis, while kit PM can recover both DNA and RNA to allow analysis of RNA viruses.
Using first-pass meconium samples, the present study aimed to compare these four DNA extraction methods to assess bacterial DNA recovery, removal of PCR inhibitors, and resulting bacterial community structures in order to define the optimal extraction method for use in meconium microbiome studies.

Methods

Sample collection

First-pass meconium was collected from five infants born by elective caesarean section to healthy mothers at King Edward Memorial Hospital, Subiaco, Western Australia with the approval of the Human Research Ethics Committee of the Western Australian Department of Health’s Women and Newborns Health Service (2015026EW). All samples were passed within 12 h of birth (mean = 6.4 h) and processed within an hour of being passed. Whole nappies were removed from infants by gloved midwives, de-identified and placed in sterile transport bags. Samples were then taken from the nappies in a level two biosafety cabinet using aseptic techniques. For each meconium sample, five aliquots of 200 ± 3 mg were taken and stored at -20°C until extraction (< 1 week).

To limit the possibility of external contamination from the nappy or infant’s skin, the outer layer of the meconium was removed using a sterile scalpel. An inner portion of meconium was then retrieved using a sterile syringe, then immediately distributed into PCR safe tubes for extraction.

Tween-80 Treatment

The tar-like consistency of meconium does not lend itself easily to DNA extraction. Meconium can block filters in spin column extractions and cause bead carryover into the eluate in magnetic bead-based extractions. After several failed extraction attempts, we identified a published method of meconium solubilisation using a 10% Tween-80 solution [223]. Samples were mixed with 1 mL sterile 10% Tween 80 to achieve liquefaction, then vortexed horizontally for 20 minutes (MoBio Vortex-Genie 2, speed setting 7) and centrifuged at 40,000 X g for 6 minutes at 4°C. The supernatant was discarded and the pellet resuspended in 1 mL UltraPure water. The samples were again
centrifuged at 40,000 X g for 6 minutes at 4°C. The supernatant was again discarded and the pellet was immediately processed with the appropriate extraction kit.

**Extractions**

One 200 mg aliquot of meconium from each infant was processed with each extraction method. The extraction kits used are described in Table 3.1. Extractions with kit PM were processed on the King Fisher Duo platform. All extractions were performed according to the manufacturer’s instructions, with the exception of kit QS, for which a lysing temperature of 90°C was used instead of 70°C, as recommended by the manufacturer to process difficult to lyse samples. One aliquot of each meconium sample was processed with kit PM following the manufacturer’s instructions, and a second set of aliquots was processed with a double inhibitor removal (IRT) step (PM2) for all but one sample, for which there was insufficient remaining sample. All samples were eluted in 100 µL of UltraPure water. An extraction control consisting of 250 µL of sterile DNA-free water was used for each kit.

<table>
<thead>
<tr>
<th>Extracion Kit</th>
<th>Manufacturer</th>
<th>Abbreviation</th>
<th>Principle</th>
<th>Bead beating component?</th>
<th>Cost per sample (USD)*</th>
<th>Completion time (hours)*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stool Mini</td>
<td>Qiagen</td>
<td>QS</td>
<td>Spin column</td>
<td>No</td>
<td>$4.58</td>
<td>2.25</td>
</tr>
<tr>
<td>Microbiome</td>
<td>Qiagen</td>
<td>QM</td>
<td>Spin column</td>
<td>Yes</td>
<td>$10.67</td>
<td>4.75</td>
</tr>
<tr>
<td>Power Soil</td>
<td>MoBio</td>
<td>PS</td>
<td>Spin column</td>
<td>Yes</td>
<td>$5.48</td>
<td>1.75</td>
</tr>
<tr>
<td>Power Microbiome</td>
<td>MoBio</td>
<td>PM</td>
<td>Magnetic beads</td>
<td>Yes</td>
<td>$5.18</td>
<td>2.50</td>
</tr>
<tr>
<td>Power Microbiome double IRT</td>
<td>MoBio</td>
<td>PM2</td>
<td>Magnetic beads</td>
<td>Yes</td>
<td>$5.18</td>
<td>2.75</td>
</tr>
</tbody>
</table>

*Cost and completion time based on processing 5 samples and 1 extraction control per batch. Completion time includes time taken to pre-treat samples with Tween-80.


**DNA yield**

DNA yield was assessed using the Qubit® dsDNA HS Assay kit with a Qubit® 2.0 fluorometer. The limit of detection was 10 pg/µL.

**Quantification of human DNA in meconium samples**

Levels of human DNA present in each meconium sample were assessed via qPCR for the human β globin gene, as previously described [224]. A standard curve was constructed using EpiTech control human DNA (Qiagen) and PCR was carried out in 20 µL reactions containing 5 µL of template or water (negative template control), 1X TaqMan Fast Advanced Master Mix (Applied Biosystems), 0.1 µM each of the forward (5’-GGGCAACGTGCTGGTCTG-3’) and reverse (5’-AGGCAGCCTGCACTGGT-3’) primers, 0.25 µM of probe (5’-FAM-CTGGCCCATCCTTTGGGAAAGAA-TAMRA-3’), and 4.2 µL of water. The PCR amplification program consisted of an initial heating step of 95°C for 20 seconds, followed by 40 cycles of 95°C for 1 second and 60°C for 20 seconds. PCR reactions were performed on a ViiA7 Real-Time PCR System (Life Technologies). All samples and controls were run in duplicate.

**16S rRNA gene qPCR**

Real-time PCR was performed to compare relative levels of bacterial DNA recovery between extraction methods. The V6 region of the 16S rRNA gene was amplified as previously described [225] in 20 µL reactions containing 5 µL of template or water (negative template control), 1X TaqMan Fast Advanced Master Mix (Applied Biosystems), 0.1 µM each of the forward (5’-TGGAGCATGTGGTTAATTCGA-3’) and reverse (5’-TGCGGGACTTAACCCAACA-3’) primers, 0.25 µM of probe (5’-FAM-CACGAGCTGACGACCATGCA-BHQ1-3’), and 4.2 µL of water. All samples and controls were run in duplicate.

**Inhibitor assessment**

In order to quantify the effect of PCR inhibitors present in meconium, undiluted, purified DNA from each kit was spiked with 0.1 ng of purified *Streptococcus agalactiae* (Group B Streptococcus - GBS) DNA. By comparing the amount of GBS
DNA recovered from spiked meconium samples to the amount of GBS DNA recovered from spiked extraction controls we were able to assess the presence of PCR inhibitors remaining after extraction with each kit. Routine culture-based testing for GBS in the vaginal tracts of each participating mother were negative. In addition, all meconium samples were confirmed as negative for GBS DNA by molecular screening (GBS targeted qPCR, as described below) prior to the spiking experiments.

5 µL of extracted meconium DNA was spiked with 0.1 ng of pure GBS DNA. This eluate + GBS DNA mix was then used as the template for qPCR. GBS DNA levels were quantified using the dltS primer/probe set as previously described [226]. PCR was carried out in 20 µL reactions containing 5 µL of template or water (negative template control), 1X TaqMan Fast Advanced Master Mix (Applied Biosystems), 0.1 µM each of the forward and reverse primers, 0.25 µM of probe, and 4.2 µL of water. PCR conditions were as described above. A standard curve was used to quantify the percent recovery of GBS DNA from each sample. All samples and controls were run in duplicate.

**Endpoint PCR**

Endpoint PCR was performed to amplify the V3-V4 region of the 16S rRNA gene for sequencing. The primers used were 341F (5’-CCTACGGGNGGCWGCAG-3’) and 785R (5’-GACTACHVGGGTATCTAATCC-3’), previously validated as providing optimal coverage of the domain Bacteria for a 400-1000 bp amplicon [227]. PCR was carried out in 50 µL reactions containing 5 µL of template or water (negative template control), 1X 360 PCR buffer (ABI), 2 mM MgCl₂, 200 µM dNTPs, 1.25U of Taq, 0.5 µM each of the forward and reverse primers, and 29.25 µL of water. The PCR amplification program consisted of an initial heating step at 94°C for 3 minutes; 40 cycles of 95°C for 30 seconds, 55°C for 30 seconds, and 72°C for 1 minute; and a final extension step of 72°C for 7 minutes. PCR reactions were performed on an Applied Biosystems Veriti Thermal Cycler. PCR products were visualised on a QIAxcel automated electrophoresis system using a DNA high resolution gel cartridge (run parameters 0M500) to confirm the presence and size of amplicons.
**Ion Torrent sequencing**

For next generation sequencing (NGS) library preparation, the PCR products were purified using the Agencourt AMPure XP Reagent (Beckman Coulter) following the manufacturer’s protocol and re-suspended in 20 µL of Low TE buffer (0.1 mM EDTA, 10 mM Tris-HCl pH 8). The purified amplicons were quantified using the Qubit Fluorometer 2.0 and Qubit dsDNA Broad Range Assay Kit (Thermo Fisher Scientific) according to the manufacturer’s protocol. To enable sequencing adaptor and sample indexing barcode ligation, the purified PCR products (100 ng) were first blunt-ended using the End Repair Enzyme Mix (Thermo Fisher Scientific) according to the manufacturer’s protocol. The Ion P1 Adaptor and Ion Xpress Barcodes 1-26 were ligated to the amplicons using the DNA Ligase Mix (Thermo Fisher Scientific) according to the manufacturer’s protocol. The adaptor-ligated libraries were purified using the Agencourt AMPure XP Reagent, re-suspended in 20 µL of Low TE Buffer and amplified using the Platinum PCR ToughMix High Fidelity and Library Amplification Primer Mix. The thermal cycling conditions consisted of an initial denaturation at 95°C for 5 minutes, followed by 5 cycles of 95°C for 15 seconds, 58°C for 15 seconds and 70°C for 1 minute. The libraries were again purified using the Agencourt AMPure XP Reagent and quantified using the Qubit dsDNA HS Assay Kit. Each library was adjusted to 100 pM in Low TE Buffer and combined in an equimolar ratio to ensure equal representation of each barcoded library in the sequencing reaction.

Automated template preparation using isothermal amplification technology and chip loading was performed using the Ion 520 & 530 ExT Kit on the Ion Chef System (Thermo Fisher Scientific). A 50 µL aliquot of the 100 pM pooled library was added to the Ion S5 ExT Reagents cartridge for templating onto Ion Sphere Particles (ISPs) and loading into an Ion 520 Chip. The loaded Ion 520 Chip was sequenced for 1300 flows using the Ion S5 ExT Sequencing Kit on an Ion S5 Sequencer with Torrent Suite Software version 5.2.2 using Default Calibration (Thermo Fisher Scientific).

**Sequencing data processing**

Sequences generated from Ion Torrent sequencing of 16S rRNA gene PCR amplicons were analysed using default setting on the open-source software Quantitative Insights
into Microbial Ecology (QIIME) version 1.9.1 [228]. Chimeras, low quality reads (Q<25), and reads of <400 bp in length were removed. Reads were assigned to operational taxonomic units using the `pick_open_reference_otus.py` command with default parameters using the UCLUST method [229]. OTUs occurring only once (singletons) or that failed to align using PyNAST were removed [230]. Relative diversity analyses were generated using the command `core_diversity_analyses.py`. Raw sequence reads have been submitted to the Sequence Read Achieve (accession number SRP128962). Phylum and genus level abundance plots were created using MicrobiomeAnalyst [231].

**Results & discussion**

**DNA yield**

DNA yield varied between meconium samples, but was even more variable according to extraction method (Figure 3.1). Kit PM produced markedly (~10-fold) higher mean DNA yield (80 ng/g meconium) compared to the other kits (QS = 5 ng/g meconium, QM = 5.15 ng/g meconium, PS = 4.85 ng/g meconium), suggesting a superior extraction efficiency. This result is particularly interesting as kit QM isolates only prokaryotic DNA, while kit PM isolates all DNA (i.e. bacterial and human). This may suggest that kit PM recovers a high quantity of human DNA from meconium samples. However, it may also indicate that kit PM is simply able to recover more DNA than kit QM in general. All negative extraction controls yielded DNA below the limit of detection (10 pg/µL, data not shown). Our extractions yielded significantly less total DNA than those reported in meconium by Wampach *et al.* (mean: 200 ng/g) [13] who employed an unspecified pre-processing step, followed by a modified kit PS protocol.
FIGURE 3.1: DNA quantitation (ng/g meconium) for each meconium sample using various extraction methods.

Quantification of human DNA in meconium samples

To assess the extent to which contaminating host DNA influenced the results from the quantification of total DNA, we performed a qPCR for human β-globin DNA. Levels of human DNA were very low, below the limit of detection (0.5 pg/µL) in several samples (1/5 from kit QS, 3/5 for kit QM, 5/5 for kit PS, and 2/5 for kit PM). Human DNA made up 15.8% of total extracted DNA from kit QS, 4.8% of total extracted DNA from kit QM, and 1.3% of total extracted DNA from kit PM. All extraction controls were below the limit of detection. Human DNA thus makes up a minor proportion of the total DNA content of meconium, which suggests that the high levels of total DNA seen in eluates from kit PM reflect greater microbial DNA extraction efficiency than the other kits, not human DNA contamination.
Contaminating bacterial DNA is ubiquitous in DNA extraction kits and other laboratory reagents [74]. Such contamination can be a major confounding factor in metagenomic studies of low-biomass samples such as meconium. For this reason, a negative extraction control was processed alongside our meconium samples to provide a point of reference for bacterial DNA contamination. The extraction controls had DNA levels below the limit of detection of the Qubit high sensitivity DNA quantitation assay (10 pg/µL). Mean cycle threshold (Ct) values for each negative extraction control were as follows: kit QS, 32.9; kit QM, 31.5; kit PS, 33.9; kit PM, 35.6; kit PM2, 34.3 (Figure 3.2, black bars). Kit PM returned the highest Ct value for its negative extraction control, suggesting the lowest amount of DNA contamination in the kit components. Two negative PCR controls were run alongside all samples and gave a mean Ct result of 33.6 ± 0.3 (data not shown). With kits QS and PS, sample extracts contained more DNA (lower Ct values) than negative extraction controls; this was not the case for samples processed with kits QM and PM. In the case of kits QS and QM, extraction control Ct values were lower than those of the negative PCR control (33.6), suggesting that these kits contributed a small amount of bacterial DNA contamination.

The greatest difference between mean Ct values of the negative extraction controls and samples was seen in kit PS, with a 1.7 cycle difference. Only 2/5 samples extracted with kit QM and only 3/5 samples extracted with kit PM yielded Ct values below the negative extraction control. (Figure 3.2). Therefore, to assess the impact of PCR inhibitors in the amplification and detection of DNA with kit PM, a second extraction was performed using a double inhibitor removal treatment (IRT) step. This extraction method (PM2) did not appreciably lower the Ct value for the same sample set, suggesting that PCR inhibition remained a significant issue.
Inhibitor assessment

Despite the use of PCR inhibitor removal steps, PCR inhibitors remained after processing with each of the four tested extraction kits. The least evidence of PCR inhibition was found in kit QM (median 9.0% inhibition), followed by kit PS (median 11.2% inhibition), and kit QS (median 12.9% inhibition) (Figure 3.3). Kit PM performed poorly in this regard, with a median 62.1% inhibition. Addition of the double inhibitor removal step had modest impact, reducing the level of inhibition to a median 43.5% (PM2). However, this did not translate to improved qPCR results (Figure 3.2).

To further investigate the high levels of PCR inhibition associated with kit PM, we normalised DNA levels in each sample extracted with kit PM2 to the average levels of DNA in eluates produced with the other 3 kits (5 ng/g meconium). After normalisation, we observed no PCR inhibition from PM2 eluates, suggesting that levels of inhibition are relative to levels of extracted DNA. Alternatively the inhibitors may have been diluted to an insignificant level.

Given the qPCR results and those of our DNA quantification, we concluded that kit PM achieves the highest yield of DNA but suffers from the greatest degree of PCR inhibition. However, this inhibition appears to be directly related to the high DNA.
yields and was completely resolved following dilution of DNA in line with levels extracted from the other three kits. It is possible that some PCR inhibitors in meconium are similarly charged to DNA, and since kit PM is magnetic-bead based, this would result in concurrent transfer of DNA and any remaining inhibitors into the eluates. We hypothesise that there is likely to be an optimal DNA dilution ratio that still maintains minimal PCR inhibition with kit PM, however, this may also be sample-specific, meaning that construction of a DNA dilution series and subsequent inhibitor assay screening could be a necessary step in DNA extraction protocols for meconium prior to downstream analyses with this kit. This will be explored in additional studies. Kit PM also has the advantage in that it is able to extract both DNA and RNA, allowing users to analyse the bacterial and viral component of the meconium microbiome. Kit PS also performed well, with low levels of inhibition from undiluted DNA as evident in 16S rRNA gene qPCR and GBS spiking experiments, however, the overall DNA yield from this kit were 16-fold less than that of kit PM.

FIGURE 3.3: Percent recovery of GBS DNA with meconium extractions generated from various extraction kits. Data are mean ± SD.
16S rRNA gene sequencing

16S rRNA gene sequencing analysis showed that the meconium microbiome was dominated by sequences affiliated with bacteria of the Proteobacteria and Firmicutes phyla, as previously reported in numerous studies [210]. However, in the present study a high level of inter-kit variation in meconium microbiome structure was observed (Figures 3.4 and 3.5). In particular, amplicons generated from DNA extracted with kit QM consisted largely of sequences affiliated with Firmicutes (with a total of 47.4% of reads belonging to this phylum), while the other kits were mainly affiliated with Proteobacteria (Figure 3.4).

The microbial profiles generated after extraction with kits QS, PS and PM were similar, with QS-extracted meconium DNA profiles dominated by sequences affiliated with unknown Comamonadaceae (23.8%), Staphylococcus sp. (15.5%), unknown Enterobacteriaceae (14.7%), and Sphingomonas sp. (9.1%); PS-extracted meconium DNA was dominated by unknown Comamonadaceae (23.5%), Staphylococcus sp. (18.3%), unknown Enterobacteriaceae (17.9%), and Sphingomonas sp. (9.1%); while PM-extracted meconium DNA was dominated by unknown Enterobacteriaceae (24.4%), unknown Comamonadaceae (22.9%), Staphylococcus sp. (11.1%), and Sphingomonas sp. (8.5%). For kit QM, however, the microbial profile was dominated by sequences affiliated with Enterococcus sp. (25.1%), Staphylococcus sp. (15.8%), unknown Comamonadaceae (15.1%), and Brachybacterium sp. (7.1%) (Figure 3.5).

The major differences in microbial profiles between kit QM and the other kits may be due to the inclusion of a eukaryotic DNA removal step in kit QM. Prior to bacterial cell lysis, host cells are selectively lysed and DNA is enzymatically degraded. It is possible that some bacterial cells may be lysed during this step, for instance those attached to human cells, thus changing the bacterial community structure. However, amplicons generated from kit QM DNA produced the highest number of unique sequences at the genus level (14, compared to 8 from PS, 2 from PM, and 1 from QS). Thus, it seems unlikely that a large diversity of bacteria is lost in this step. Although the precise host DNA removal methods used in the kit are proprietary and not disclosed, previous studies have demonstrated that removal of human DNA through selective lysis of eukaryotic cells is not 100% efficient, and results in some loss of bacterial DNA [232].
FIGURE 3.4: Relative abundance of OTUs in meconium samples (n = 5) after extraction with various kits at phylum level.
FIGURE 3.5: Relative abundance of OTUs in meconium samples (n = 5) after extraction with various kits at genus level.

It is difficult to compare our results to the “true” meconium microbiome, as it is a poorly studied substance. Instead we tested the reproducibility of each extraction kit per-patient by quantifying its ability to recover the aggregate microbiome for each patient. All OTUs recovered from a single patient across all kits were pooled, and each kit was scored by its ability to recover this pooled microbiome for each patient. Kit QM recovered the highest percentage of OTUs per patient, 44-79%. Kit QS was able to recover 20-47%, kit PS recovered 33-55%, and kit PM recovered 12-47%. Using this rationale, it appears that kit QM is best able to extract the aggregate meconium
microbiome. To better assess the ability of these kits to extract the “true” microbiome of meconium, a mock community may be added to these samples.

A number of OTUs were detected in our negative extraction controls (14 from kit QS, 38 from kit QM, 17 from kit PS, 23 from kit PM) and in our negative PCR controls (21) (Table 3.2). It is now well established that negative extraction and PCR controls contain trace amounts of microbial DNA [74, 233], and that contamination from extraction kits and laboratory reagents is a major confounding issue when working with low biomass samples such as meconium [234]. 51 OTU sequences were found only in meconium samples, not in negative extraction controls or PCR controls (Table 3.3). Thus, we can say with some certainty that the source of these bacterial sequences was meconium.

This study provides strong evidence that choice of DNA extraction kit impacts upon 16S rRNA gene microbial profiles generated from first-pass meconium samples. This data is in line with previous studies that have shown a high level of variation in microbiome community structure following DNA extraction with different kits [235-239]. This reinforces the need to develop a standardised, validated meconium extraction protocol so that results may be compared between studies. We have also shown that reagent and kit contamination can confound microbiome studies on meconium samples (Figure 3.6). Sequences generated from negative extraction and PCR controls must be taken into account when analysing meconium microbiome profiles and data must be interpreted cautiously.

Given our data, we recommend the use of kits QM and PM for microbiome analysis of meconium. However, we caution that given the variation in OTUs recovered between kit QM and the other kits tested, results gained from use of kit QM cannot be compared to those produced with other kits in other studies.
FIGURE 3.6: Comparison of relative abundance of the 10 most abundant OTUs in meconium samples with extraction controls and PCR controls from A) kit QS, B) kit QM, C) kit PS, and D) kit PM. Note that abundances are relative and not absolute, so a direct comparison of abundance cannot be made between meconium samples, PCR controls and extraction controls.
Summary

We compared four commonly used DNA extraction methods to assess their ability to extract DNA, overcome PCR inhibitors and analyse bacterial DNA from meconium. Our results indicate that kit PM is best able to extract microbial DNA from meconium; however, eluates require dilution to remove PCR inhibitors. We have also demonstrated a high level of variation in microbiome community structure after extraction with different kits, and the importance of controlling for external DNA contamination. Eluates generated with kit QM differed significantly from those generated with the other kits in terms of the dominant phyla and genera. The other three kits were consistent in terms of dominant taxa, but differed significantly in terms of low abundance OTUs.

Our results indicate that there are very low levels of human DNA in meconium relative to levels of microbial DNA and highlight the need to establish a meconium-specific sampling/extraction protocol for microbiome studies on the fetal gut. Again, we emphasise the importance of negative extraction controls for work in low biomass samples such as meconium.
TABLE 3.2: Summary of OTUs detected in negative extraction controls (n = 1 per kit) and negative PCR controls (n = 2).

<table>
<thead>
<tr>
<th>Kit QS</th>
<th>OTU</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Burkholderiales;f_Comamonadaceae;g__</td>
<td>0.450292</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Alphaproteobacteria;o_Sphingomonadales;f_Sphingomonadaceae;g__Sphingomonas</td>
<td>0.280298</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Alphaproteobacteria;o_Rhizobiales;f_Methylobacteriaceae;g__</td>
<td>0.149029</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Burkholderiales;f_Comamonadaceae;g__Comamonas</td>
<td>0.071184</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_TM7;Other;Other;Other;Other</td>
<td>0.033878</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Gammaphyla;g__Sphingomonas</td>
<td>0.005041</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Burkholderiales;f_Oxalobacteriaceae;g__Ralstonia</td>
<td>0.003831</td>
</tr>
<tr>
<td></td>
<td>Unassigned;Other;Other;Other;Other</td>
<td>0.001613</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Firmicutes;c_Bacilli;o_Bacillales;f_Staphylococcaceae;g__Staphylococcus</td>
<td>0.00121</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Alphaproteobacteria;o_Sphingomonadales;f_Sphingomonadaceae;g__</td>
<td>0.000605</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Rhodocyclaceae;f_Rhodocyclaceae;g__</td>
<td>0.000605</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g__Streptococcus</td>
<td>0.000202</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Gammaphyla;g__Enterobacteriaceae;g__Enterobacteriaceae</td>
<td>0.000202</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Acinobacteria;c_Acinobacteriaceae;o_Acinobacteriaceae;f_Acinobacteriaceae;g__Acinobacter</td>
<td>0.202392</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Acinobacteria;c_Acinobacteriaceae;o_Acinobacteriaceae;f_Dermabacteriaceae;g__Brachybacterium</td>
<td>0.137781</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Acinobacteria;c_Acinobacteriaceae;o_Acinobacteriaceae;f_Brevibacteriaceae;g__Brevibacterium</td>
<td>0.091815</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Acinobacteria;c_Acinobacteriaceae;o_Acinobacteriaceae;f_Corynebacteriaceae;g__Corynebacterium</td>
<td>0.079386</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Enterococcaceae;g__Enterococcus</td>
<td>0.064259</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Firmicutes;c_Bacilli;o_Bacillales;f_Pseudomonas;g__Pseudomonas</td>
<td>0.058665</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Burkholderiales;f_Comamonadaceae;g__</td>
<td>0.057575</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Acinobacteriaceae;o_Acinobacteriaceae;f_Methylobacteriaceae;g__Methylobacterium</td>
<td>0.049367</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Lactobacillaceae;g__Lactobacillus</td>
<td>0.031191</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Acinobacteria;c_Acinobacteriaceae;o_Acinobacteriaceae;f_Propionibacteriaceae;g__Propionibacterium</td>
<td>0.026853</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Acinobacteria;c_Acinobacteriaceae;o_Acinobacteriaceae;f_Micrococciaceae;g__Micrococcus</td>
<td>0.022397</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Aerococcaceae;g__Aerococcus</td>
<td>0.021459</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Streptococcaceae;g__Streptococcus</td>
<td>0.015713</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Gammaphyla;g__Enterobacteriaceae;g__Enterobacteriaceae</td>
<td>0.014544</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Proteobacteria;c_Gammaphyla;g__Enterobacteriaceae;g__Enterobacteriaceae</td>
<td>0.013368</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Firmicutes;c_Bacilli;o_Lactobacillales;f_Aerococcaceae;g__Facklamia</td>
<td>0.011257</td>
</tr>
<tr>
<td></td>
<td>k_Bacteria;p_Acinobacteria;c_Acinobacteriaceae;o_Acinobacteriaceae;f_Kineosporiaceae;g__</td>
<td>0.011143</td>
</tr>
<tr>
<td>Kingdom</td>
<td>Phylum</td>
<td>Class</td>
</tr>
<tr>
<td>-----------</td>
<td>----------------</td>
<td>-----------------</td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Firmicutes</strong></td>
<td><strong>Bacilli</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Alphaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Alphaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Deltaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Alphaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Actinobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Gammaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Gammaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Bacteria</strong></td>
<td><strong>Firmicutes</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Betaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Actinobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Alphaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Betaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Cyanobacteria</strong></td>
<td><strong>4C0d-2</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Bacteroidetes</strong></td>
<td><strong>Flavobacteriia</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Bacteroidetes</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Alphaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Bacteroidetes</strong></td>
<td><strong>Flavobacteriia</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Alphaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Betaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Bacteroidetes</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Alphaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Gammaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Betaproteobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Proteobacteria</strong></td>
<td><strong>Actinobacteria</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Bacteroidetes</strong></td>
<td><strong>Bacteroidales</strong></td>
</tr>
<tr>
<td><strong>Bacteria</strong></td>
<td><strong>Bacteroidetes</strong></td>
<td><strong>Bacteroidales</strong></td>
</tr>
</tbody>
</table>

**Kit PS**

<table>
<thead>
<tr>
<th>OTU</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>k_Bacteria:p Proteobacteria:c Betaproteobacteria:o Burkholderiales:f Comamonadaceae:g Comamonas</td>
<td>0.347242</td>
</tr>
<tr>
<td>k_Bacteria:p Proteobacteria:c Betaproteobacteria:o Burkholderiales:f Comamonadaceae:g Delftia</td>
<td>0.151218</td>
</tr>
<tr>
<td>k_Bacteria:p Proteobacteria:c Alphaproteobacteria:o Sphingomonadales:f Sphingomonadaceae:g Sphingomonas</td>
<td>0.127975</td>
</tr>
<tr>
<td>k_Bacteria:p Proteobacteria:c Betaproteobacteria:o Burkholderiales:f Comamonadaceae:g Comamonas</td>
<td>0.105853</td>
</tr>
<tr>
<td>k_Bacteria:p Bacteroidetes:c Flavobacteriia:o Flavobacteriales:f Flavobacteriaceae:g Capnocytophaga</td>
<td>0.079250</td>
</tr>
<tr>
<td>k_Bacteria:p Proteobacteria:c Alphaproteobacteria:o Sphingomonadales:f Sphingomonadaceae:g Sphingomonas</td>
<td>0.055447</td>
</tr>
<tr>
<td>k_Bacteria:p Bacteroidetes:c Bacteroidales:f Prevotellaceae:g Prevotella</td>
<td>0.032484</td>
</tr>
<tr>
<td>k_Bacteria:p Proteobacteria:c Gammaproteobacteria:o Pseudomonadales:f Moraxellaceae:g Acinetobacter</td>
<td>0.031364</td>
</tr>
<tr>
<td>k_Bacteria:p Proteobacteria:c Gammaproteobacteria:o Pseudomonadales:f Moraxellaceae:g Acinetobacter</td>
<td>0.018202</td>
</tr>
<tr>
<td>k_Bacteria:p Proteobacteria:c Gammaproteobacteria:o Pasteurellales:f Pasteurellaceae:g Actinobacillus</td>
<td>0.013722</td>
</tr>
<tr>
<td>k_Bacteria:p Actinobacteria:c Actinobacteria:o Actinomycetales:f Propionibacteriaceae:g Propionibacterium</td>
<td>0.013162</td>
</tr>
<tr>
<td>k_Bacteria:p Firmicutes:c Clostridia:o Clostridiales:f Veillonellaceae:g Veillonella</td>
<td>0.011201</td>
</tr>
<tr>
<td>k_Bacteria:p Bacteroidetes:c Bacteroidales:f Porphyromonadaceae:g Porphyromonas</td>
<td>0.007001</td>
</tr>
<tr>
<td>k_Bacteria:p Firmicutes:c Bacilli:o Bacillales:f Staphylococcaceae:g Staphylococcus</td>
<td>0.004201</td>
</tr>
</tbody>
</table>
### Kit PM

<table>
<thead>
<tr>
<th>OTU</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Burkholderiales;f_Comamonadaceae;g_Other</td>
<td>0.00084</td>
</tr>
<tr>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Burkholderiales;f_Comamonadaceae;g_Other</td>
<td>0.00056</td>
</tr>
<tr>
<td>k_Bacteria;p_Actinobacteria;c_Actinobacteria;o_Actinomycetales;f_Brevibacteriaceae;g_Brevibacterium</td>
<td>0.00028</td>
</tr>
</tbody>
</table>

### PCR control

<table>
<thead>
<tr>
<th>OTU</th>
<th>Relative abundance</th>
</tr>
</thead>
<tbody>
<tr>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Burkholderiales;f_Comamonadaceae;g_Other</td>
<td>0.555856</td>
</tr>
<tr>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Rhizobiales;f_Methylobacteriaceae;g_Methylobacterium</td>
<td>0.154918</td>
</tr>
<tr>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Rhizobiales;f_Methylobacteriaceae;g_Other</td>
<td>0.000213</td>
</tr>
<tr>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Rhizobiales;f_Methylobacteriaceae;g_Other</td>
<td>0.051636</td>
</tr>
<tr>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Rhizobiales;f_Oxalobacteraceae;g_Ralstonia</td>
<td>0.011103</td>
</tr>
<tr>
<td>k_Bacteria;p_Proteobacteria;c_Betaproteobacteria;o_Burkholderiales;f_Comamonadaceae;g_Other</td>
<td>0.002081</td>
</tr>
<tr>
<td>Kingdom</td>
<td>Phylum</td>
</tr>
<tr>
<td>---------</td>
<td>--------------</td>
</tr>
<tr>
<td>Bacteria</td>
<td>Proteobacteria</td>
</tr>
<tr>
<td></td>
<td>Firmicutes</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Proteobacteria</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria</td>
</tr>
<tr>
<td></td>
<td>Firmicutes</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Proteobacteria</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria</td>
</tr>
<tr>
<td></td>
<td>Actinobacteria</td>
</tr>
<tr>
<td></td>
<td>Actinobacteria</td>
</tr>
<tr>
<td></td>
<td>Bacteroidetes</td>
</tr>
<tr>
<td></td>
<td>Firmicutes</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria</td>
</tr>
<tr>
<td></td>
<td>Proteobacteria</td>
</tr>
</tbody>
</table>
**TABLE 3.3:** Summary of OTUs detected in meconium samples only, not in negative extraction controls or negative PCR controls.

<table>
<thead>
<tr>
<th>Phyla</th>
<th>Class</th>
<th>Order</th>
<th>Family</th>
<th>Genus</th>
<th>Recovered by kit</th>
</tr>
</thead>
<tbody>
<tr>
<td>Actinobacteria</td>
<td>Actinomycetales</td>
<td>Unknown</td>
<td>Unknown</td>
<td>QM, PS</td>
<td></td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>Actinomycetales</td>
<td>Microbacteriaceae</td>
<td>Unknown</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>Actinomycetales</td>
<td>Micrococcaceae</td>
<td>Arthrobacter</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>Actinomycetales</td>
<td>Micrococcaceae</td>
<td>Kocuria</td>
<td>QM, PM</td>
<td></td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>Actinomycetales</td>
<td>Nocardiaceae</td>
<td>Rhodococcus</td>
<td>QM, QM, PS</td>
<td></td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>Actinomycetales</td>
<td>Pseudonocardia</td>
<td>Pseudonocardia</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Actinobacteria</td>
<td>Actinomycetales</td>
<td>Streptomycescaceae</td>
<td>Streptomyces</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Bacteroidetes</td>
<td>Flavobacteria</td>
<td>Flavobacterales</td>
<td>Weeksellaceae</td>
<td>Cloacibacterium</td>
<td>PM</td>
</tr>
<tr>
<td>Saprospirae</td>
<td>Saprospirales</td>
<td>Chitinophagaceae</td>
<td>Sediminibacterium</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Cyanobacteria</td>
<td>Chloroplast</td>
<td>Streptophyta</td>
<td>Unknown</td>
<td>Unknown</td>
<td>QM, QM</td>
</tr>
<tr>
<td>Bacilli</td>
<td>Bacillales</td>
<td>Bacillaceae</td>
<td>Geobacillus</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Bacilli</td>
<td>Bacillales</td>
<td>Paenibacillaceae</td>
<td>Paenibacillus</td>
<td>QM, QM</td>
<td></td>
</tr>
<tr>
<td>Bacilli</td>
<td>Bacillales</td>
<td>Planococcaceae</td>
<td>Unknown</td>
<td>QM, PS</td>
<td></td>
</tr>
<tr>
<td>Bacilli</td>
<td>Gemellales</td>
<td>Gemellaceae</td>
<td>Unknown</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Bacilli</td>
<td>Lactobacillales</td>
<td>Unknown</td>
<td>Unknown</td>
<td>QM, QM, PM</td>
<td></td>
</tr>
<tr>
<td>Bacilli</td>
<td>Lactobacillales</td>
<td>Aerococcaceae</td>
<td>Marinilactibacillus</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Bacilli</td>
<td>Lactobacillales</td>
<td>Lactobacillaceae</td>
<td>Unknown</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Bacilli</td>
<td>Lactobacillales</td>
<td>Streptococcaceae</td>
<td>Lactococcus</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Clostridia</td>
<td>Clostridiales</td>
<td>Clostridiaceae</td>
<td>Clostridium</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Clostridia</td>
<td>Clostridiales</td>
<td>Ruminococcaceae</td>
<td>Unknown</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Erysiipelotrichi</td>
<td>Erysiipelotrichales</td>
<td>Erysiipelotrichaceae</td>
<td>Catenibacterium</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Gemmatimonadetes</td>
<td>Gemmatimonadetes</td>
<td>N1423VL</td>
<td>Unknown</td>
<td>Unknown</td>
<td>QM</td>
</tr>
<tr>
<td>Nitrospira</td>
<td>Nitrospira</td>
<td>Nitrospirales</td>
<td>Nitrospiraceae</td>
<td>Nitrospira</td>
<td>QM</td>
</tr>
<tr>
<td>Planctomycetes</td>
<td>Planctomycetia</td>
<td>Planctomycetaceae</td>
<td>Planctomyces</td>
<td>Planctomyces</td>
<td>QM</td>
</tr>
<tr>
<td>Proteobacteria</td>
<td>Alphaproteobacteria</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>QM, QM</td>
</tr>
<tr>
<td>Alphaproteobacteria</td>
<td>Caulobacterales</td>
<td>Caulobacteraceae</td>
<td>Brevundimonas</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Alphaproteobacteria</td>
<td>Rhodobacterales</td>
<td>Rhodobacteraceae</td>
<td>Paracoccus</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Alphaproteobacteria</td>
<td>Rhodobacterales</td>
<td>Rhodobacteraceae</td>
<td>Rhodobacter</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Alphaproteobacteria</td>
<td>Rhodospirillales</td>
<td>Acetobacteraceae</td>
<td>Unknown</td>
<td>QM, QM</td>
<td></td>
</tr>
<tr>
<td>Alphaproteobacteria</td>
<td>Rickettsiales</td>
<td>Unknown</td>
<td>Unknown</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Proteobacteria Class</td>
<td>Order</td>
<td>Family</td>
<td>Genus</td>
<td>QS, QM, PM</td>
<td></td>
</tr>
<tr>
<td>---------------------------</td>
<td>--------------------------------</td>
<td>---------------------------------</td>
<td>---------------------</td>
<td>------------</td>
<td></td>
</tr>
<tr>
<td>Alphaproteobacteria</td>
<td>Sphingomonadales</td>
<td>Sphingomonadaceae</td>
<td>Sphingobium</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>Unknown</td>
<td>Unknown</td>
<td>Unknown</td>
<td>QM, PS, PM</td>
<td></td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>Burkholderiales</td>
<td>Unknown</td>
<td>Unknown</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>Burkholderiales</td>
<td>Acaligenaceae</td>
<td>Achromobacter</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>Neisseriales</td>
<td>Neisseriaceae</td>
<td>Unknown</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>Rhodocyclales</td>
<td>Rhodocyclaceae</td>
<td>Hydrogenophilus</td>
<td>QM</td>
<td></td>
</tr>
<tr>
<td>Betaproteobacteria</td>
<td>Rhodocyclales</td>
<td>Rhodocyclaceae</td>
<td>Zoogloea</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Deltaproteobacteria</td>
<td>Bdellovibrionales</td>
<td>Bacteriovoracaceae</td>
<td>Unknown</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>Alteromonadales</td>
<td>Chromatiaceae</td>
<td>Rheinheimera</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>Enterobacteriales</td>
<td>Enterobacteriaceae</td>
<td>Unknown</td>
<td>QM, PS, PM</td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>Enterobacteriales</td>
<td>Enterobacteriaceae</td>
<td>Citrobacter</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>Enterobacteriales</td>
<td>Enterobacteriaceae</td>
<td>Erwinia</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>Legionellas</td>
<td>Legionellaceae</td>
<td>Unknown</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>Legionellas</td>
<td>Legionellaceae</td>
<td>Legionella</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>Pseudomonadales</td>
<td>Pseudomonadaceae</td>
<td>Unknown</td>
<td>PS</td>
<td></td>
</tr>
<tr>
<td>Gammaproteobacteria</td>
<td>Xanthomonadales</td>
<td>Xanthomonadaceae</td>
<td>Unknown</td>
<td>QM, PS</td>
<td></td>
</tr>
</tbody>
</table>
Chapter 4

Identification and removal of contaminating microbial DNA from PCR reagents: impact on low biomass microbiome analyses

This chapter also addresses Aim 1 of this thesis – to refine the present methodologies for studying the fetal microbiome. Contamination is a major problem in studies of low biomass microbiomes. Studies of the fetal gut microbiome are complicated by the fact that meconium typically contains very low titres of bacteria. Low levels of contamination in laboratory reagents can therefore overshadow the true fetal gut microbiome. In this chapter the major source of reagent contamination in microbiome studies is identified and eliminated, allowing for more sensitive and robust studies of the fetal microbiome.

Abstract

Reagent-derived contamination can compromise the integrity of microbiome data, particularly in low microbial biomass samples. This contamination has largely been attributed to the “kitome” (contamination introduced by the DNA extraction kit), prior to which attention was mostly paid to potential contamination introduced by PCR reagents. In this study, we assessed the proportion to which our DNA extraction kit and PCR master mix introduce contaminating microbial DNA to bacterial microbial profiles generated by 16S rRNA gene sequencing. Utilising a commercial dsDNase treatment protocol to decontaminate the PCR master mix, we demonstrated that the vast majority of contaminating DNA was derived from the PCR master mix – the “mixome”. Importantly, this “mixome” was almost completely eliminated using the simple dsDNase treatment, resulting in a 99% reduction in contaminating bacterial reads. We suggest that dsDNase treatment of PCR reagents should be explored as a simple and effective way of reducing contamination in low-biomass microbiome studies and producing more robust and reliable data.

Background

The user-friendly workflows and affordable nature of 16S rRNA gene sequencing has given rise to the increasing popularity of microbiome studies. However, in recent years it has become apparent that many studies are plagued by entrenched methodological errors, effectively resulting in the production of large amounts of erroneous data which have been spuriously interpreted. While it is often possible to retrospectively clean up the data [240], such corrections are not regularly made, and may themselves introduce bias. Many microbiome studies fail to use appropriate negative extraction and PCR controls to allow identification of contamination during sample processing. This is particularly problematic in low biomass samples, where low level contamination may be mistaken for a true microbiome [241]. Alternatively, studies on low biomass samples or samples with high levels of PCR inhibitors that do use appropriate controls may find similar levels of bacterial DNA in their samples and controls. These samples may, therefore, be deemed sterile, despite the fact that they may contain low levels of microbial DNA. There is, therefore, a need to rigorously define and control for contamination that might be introduced during extraction and PCR amplification processes.
Previous studies have introduced the idea of the “kitome” – the ubiquitous contamination that is present in DNA extraction kits and laboratory reagents/plasticware used during the extraction process [71, 73, 74, 76, 242]. Similarly, several studies have reported that reagents used in PCR master mixes, including polymerases, buffers, primers, and molecular biology grade water, contain low level contamination [70, 72, 75, 243, 244].

A number of attempts have been made to eliminate background microbial DNA contamination originating from PCR reagents, including UV irradiation [70], restriction endonuclease digestion [70, 245, 246], DNase treatment [70, 247], ultrafiltration to remove high molecular weight DNA [248], and treatment with DNA intercalating products [72, 249]. However, for the most part these attempts have not been able to reliably and reproducibly eradicate background contamination, and in cases where they appeared to be successful, the methods used to confirm this lacked the sensitivity of contemporary next-generation sequencing. Further, several of these methods are technically complicated, unsuitable for high-throughput work, and in many cases decontamination procedures resulted in decreased PCR sensitivity [70].

A rapid, effective, reproducible method for decontamination of PCR reagents prior to amplification of universal bacterial genes is needed for incorporation into current bacterial microbiome workflows, especially those where low microbial biomass samples are being examined. Although previous studies have shown that both DNase and endonuclease treatment can be effective options, the protocols employed were far from ideal considering the PCR master mix cannot be treated with the inclusion of the polymerase as enzyme inactivation is generally conducted at temperatures that would activate even hot-start DNA polymerases. Further, PCR primers cannot be treated as DNases will also digest single-stranded DNA and endonucleases may cleave sites within the primers themselves. A potential simple solution to these problems is the use of recently described double-strand specific DNases from the northern shrimp (*Pandalus borealis*), which can be irreversibly inactivated at 65°C [250]. Champlot *et al.* reported on the effectiveness of these; however, the reported multi-strategy reagent decontamination procedure is not ideal in terms of time and cost for incorporation into high throughput 16S rRNA gene sequencing protocols [251]. These dsDNases are now
commercially available, optimised for use in molecular biology workflows and allow single-step treatment of complete PCR master mixes in 40 minutes with no impact on PCR sensitivity. The aim of the present study was to utilise a commercial dsDNase treatment protocol to decontaminate our in-house reagents prior to 16S rRNA gene amplification and next-generation sequencing of no template control samples. The goal was to determine whether microbial DNA contamination that exists in 16S rRNA gene studies truly originates from the DNA extraction kit plasticware/reagents (the “kitome”), or whether it is derived from the PCR plasticware/reagents (the “mixome”), or a combination of the two.

**Methods**

**DNA extraction**

Four blank extraction controls (no sample input) were processed using the Qiagen MagAttract PowerMicrobiome DNA/RNA kit (previously owned by MoBio) on the Kingfisher Duo platform according to the manufacturer’s instructions. Each blank control was extracted in a separate batch and run on the KF Duo separately.

**DNA quantification**

DNA yield was assessed using the Qubit® dsDNA HS Assay kit with a Qubit® 2.0 fluorometer. The limit of detection was 10 pg/µL.

**Master mix decontamination**

For each PCR, the four extraction controls (EC1-4) and two blank PCR controls (no template controls, referred to herein as NT1 & NT2) were run with and without master mix decontamination. For the decontaminated samples, PCR master mix solutions were treated with a PCR Decontamination Kit (ArcticZymes®), which consisted of a double strand-specific DNase (dsDNase) and DTT (which aids in the inactivation of the dsDNase), as per the manufacturer’s instructions. Briefly, master mix solutions (including primers, and in the case of qPCR, probes) were treated with 0.5 µL of dsDNase and 0.5 µL of DTT per 20 µL reaction (volumes were adjusted accordingly for
50 µL endpoint reactions), then incubated at 37°C for 20 minutes (dsDNase activation), followed by incubation at 60°C for 20 minutes (dsDNase inactivation). After this, decontaminated master mix solutions were cooled on ice for two minutes to prevent heat-related changes in volume.

**16S rRNA gene qPCR**

Real-time PCR was performed to compare relative levels of bacterial DNA between dsDNase-treated and untreated extraction and PCR controls. The V6 region of the 16S rRNA gene was amplified using the primer set 891F (5′-TGGAGCATGTGGTTTAATTCGA-3′) and 1033R (5′-TGCGGGACTTAAACCAACA-3′) as previously described [252]. Briefly, 20 µL reactions were set up containing 5 µL of template (blank extraction control or nuclease-free water (Integrated DNA Technologies)), 1X TaqMan Fast Advanced Master Mix (Applied Biosystems), 0.9 µM each of the forward and reverse primers, 0.25 µM of probe (5′-FAM-CACGAGCTGACGACARCCATGCA-TAMRA-3′), and 4.2 µL of water. 40 cycles of amplification were performed on the ViiA 7 Real-Time PCR System using the Taqman Fast settings. All samples were run in duplicate.

**Endpoint PCR**

PCR amplification was carried out for next generation sequencing (NGS) as previously described [252]. The V3-V4 region of the 16S rRNA gene was amplified using the primer pair 341F (5′-CCTACGGGGNGCCWGCAG-3′) and 785R (5′-GACTACHVGGGTATCTAATCC-3′) [227]. PCR was carried out in 50 µL reactions containing 15 µL of template or water (negative template control), 1X AccuStart II ToughMix (Quantabio), 0.3 µM each of the forward and reverse primers, 1.25 µL each of the ArcticZymes PCR Decontamination Kit dsDNase and DTT or water (untreated mix), and water to bring the volume to 50 µL. The PCR amplification program consisted of an initial heating step at 94°C for 3 minutes; 40 cycles of 94°C for 30 seconds, 55°C for 30 seconds, 72°C for 2 minutes; and a final extension step of 72°C for 7 minutes. PCR reactions were performed on an Applied Biosystems Veriti Thermal Cycler. PCR products were visualised on a QIAxcel automated electrophoresis system.
using a DNA high resolution gel cartridge (run parameters OM500) to confirm the presence and size of amplicons.

16S rRNA gene sequencing

PCR products were purified using the Agencourt AMPure XP Reagent (Beckman Coulter) following the manufacturer’s protocol and re-suspended in 20 µL of low TE buffer (0.1 mM EDTA, 10 mM Tris-HCl, pH 8). NGS library preparation was carried out using the Ion Plus Fragment Library Kit following the manufacturer’s protocol. Briefly, purified amplicons (100 ng) were first blunt-ended using the End Repair Enzyme Mix. The Ion P1 Adaptor and Ion Xpress Barcodes were ligated to the amplicons while simultaneous nick repair was carried out. The adaptor-ligated libraries were purified using the Agencourt AMPure XP Reagent, re-suspended in 20 µL of Low TE Buffer and amplified using the Platinum PCR ToughMix High Fidelity and Library Amplification Primer Mix. The thermal cycling conditions consisted of an initial denaturation at 95°C for 5 minutes, followed by 5 cycles of 95°C for 15 seconds, 58°C for 15 seconds and 70°C for 1 minute. The libraries were again purified using the Agencourt AMPure XP Reagent and quantified using the Qubit dsDNA HS Assay Kit. Each library was adjusted to 100 pM in low TE Buffer. Untreated samples were combined in an equimolar ratio to ensure equal representation of each barcoded library in the sequencing reaction. Untreated samples did not produce an amplicon and could therefore not be added in equimolar concentrations. Instead, these were in equal volumes (13 µL) to the final pool.

Automated template preparation using isothermal amplification technology and chip loading was performed using the Ion 520 & 530 ExT Kit on the Ion Chef System (Thermo Fisher Scientific). A 50 µL aliquot of the 100 pM pooled library was added to the Ion S5 ExT Reagents cartridge for templating onto Ion Sphere Particles (ISPs) and loading into an Ion 530 Chip. The loaded Ion 530 Chip was sequenced for 1300 flows using the Ion S5 ExT Sequencing Kit on an Ion S5 Sequencer with Torrent Suite Software version 5.2.2 using Default Calibration (Thermo Fisher Scientific).
Data processing

Demultiplexed reads were quality and length filtered, clustered, and given taxonomic assignments using GHAP v2.1. GHAP is an in-house amplicon processing pipeline developed by Paul Greenfield (CSIRO, Australia) [253] built around tools from USEarch [229] and RDP [254], combined with locally-written tools for demultiplexing and generating OTU tables.

Results & discussion

Blank extraction controls were generated from four separate DNA extractions (EC1-4). DNA concentrations in all extraction controls were determined using a high sensitivity Qubit assay, but were all below the limit of detection (10 pg/µL), suggesting that contamination introduced during extraction was minimal. Importantly, all plasticware used in this study was certified sterile, and free of nucleotides, DNA, RNA, DNases, and RNases.

To test the extent to which extraction-based and PCR-based contamination can influence qPCR assays, EC1-4 and two blank PCR controls (no template controls, NT1 & NT2) were amplified with or without master mix decontamination (PCR Decontamination Kit, ArcticZymes). PCR controls that were amplified with decontaminated master mix did not produce any signal after 40 cycles, while those amplified with untreated master mix became positive after 29.8±0.4 cycles (Table 4.1). PCR controls amplified with untreated master mix were indistinguishable from extractions controls amplified with untreated master mix, which became positive after 30.4±0.2 cycles. Treatment of the master mix with dsDNase reduced the amount of DNA detected in the extraction controls, increasing the mean Ct to 34.8±0.8 cycles, suggesting that the higher levels of DNA detected using the untreated master mix were largely attributed to contaminating microbial DNA in the PCR reagents – the “mixome”.
**TABLE 4.1:** 16S rRNA gene qPCR Ct values from blank extraction controls (EC1-4) and PCR controls (NT1-2) amplified with dsDNase-treated or untreated master mix.

<table>
<thead>
<tr>
<th></th>
<th>EC1</th>
<th>EC2</th>
<th>EC3</th>
<th>EC4</th>
<th>NT1</th>
<th>NT2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Untreated</td>
<td>30.21</td>
<td>30.27</td>
<td>30.31</td>
<td>30.66</td>
<td>29.49</td>
<td>30.09</td>
</tr>
<tr>
<td>dsDNase-treated</td>
<td>33.98</td>
<td>34.68</td>
<td>34.52</td>
<td>35.96</td>
<td>No amplification</td>
<td>No amplification</td>
</tr>
</tbody>
</table>

Extractions controls and PCR controls were amplified by endpoint PCR with and without master mix decontamination, then visualised on a QIAxcel automated electrophoresis system using a DNA high resolution gel cartridge. No amplicons could be detected from extraction controls or PCR controls that were amplified with decontaminated master mix. All extraction controls that were amplified with untreated master mix produced an amplicon of 470.5±1.7 bp with an average concentration of 0.73±0.3 ng/µL. Similarly, untreated PCR controls produced an amplicon of 471.5±0.7 bp with an average concentration of 0.56±0.05 ng/µL.

Sequencing of the blank extraction and PCR controls revealed a substantial reduction in contamination after dsDNase treatment (**Table 4.2**). Extraction controls that were amplified with decontaminated master mix produced an average of 563 reads (range: 11-2166). Similarly, blank PCR controls that were amplified with decontaminated master mix produced an average of 761 reads (range: 177-1346). In comparison, extraction controls amplified with untreated master mix produced an average of 50,917 reads (range: 47,351-56,223). PCR controls amplified with untreated master mix produced an average of 187,470 reads (range: 46,316-328,624). dsDNase treatment also reduced the diversity of OTUs observed in the controls from an average of 10.3 OTUs in the untreated controls (range: 8-12) to an average of 3.2 OTUs in the treated controls (range: 1-5).

The majority of sequences produced in extraction and PCR controls that were amplified with either dsDNase-treated or untreated master mix mapped to the Escherichia/Shigella and Pelomonas genera. The reads attributed to Escherichia/Shigella may represent
remnant DNA from the *E. coli* used to produce *Taq* DNA polymerase. *Pelomonas* sp. has previously been described as a contaminant of pure and ultra-pure waters [255, 256].

Four OTUs appeared in the treated controls but were absent in the untreated controls. These reads mapped to *Sphingomonas* sp., *Micrococcus* sp., *Cupriavidus* sp., and an unknown *Neisseriaceae*. Given that amplicons were heavily diluted prior to sequencing, it might be that some low abundance OTUs were missed by chance. Each of these four OTUs only appeared in low numbers (14-315 reads). Alternatively, it is possible that other template was preferentially amplified in the untreated controls, and that reduction of this other template in the treated controls allowed the primers to bind to the remaining DNA and be amplified.

Amplification of extraction controls with decontaminated master mix increased the Ct value of the qPCR assay by an average of 4.4 cycles (*Table 4.1*). This shift in cycle threshold was associated with a large reduction in the number of bacterial reads produced by sequencing. This suggests that the majority of the contaminants in this workflow originate from the PCR reagents, while only a small proportion of the amplification seen could be attributed to DNA that originated from the extraction kit.

While dsDNase treatment did not eradicate contamination completely, a 99% reduction in total number of contaminating sequences was observed in this study. While we have followed the manufacturer’s protocol for decontamination, it is possible that the protocol might be optimised by increasing treatment time for more effective removal of contaminating DNA.

A limitation of this study is the use of only one extraction method and one PCR master mix. We have sought to determine the extent to which the DNA extraction kit, which we have previously found to be optimal for our work, contributed to the contamination seen in our controls [252]. Contamination introduced with our extraction method was minimal. However, this may not hold true for other extraction methods, particularly spin
column-based extractions, which involve multiple un-capping and re-capping of tubes and centrifugation of open tubes in many cases. Furthermore, contamination introduced during DNA extraction and amplification varies between batches and even lot numbers of the same reagents [74, 240, 241]. It is important, therefore, for users to carry out similar tests using their own preferred reagents.

A further limitation of this study is the use of only four extraction controls and two PCR controls. The majority of sequences in our untreated controls belonged to three OTUs (Escherichia/Shigella sp., Pelomonas sp., and Sphingomonas sp.) which appeared consistently at similar levels across all controls, suggesting that they originate from the PCR mastermix. However, several OTUs identified in these controls were variably detected, and so can’t be tied definitively to the PCR mastermix. We are not the first to demonstrate the sporadic nature of microbiome contamination [62, 74]. Although we were meticulous in using sterile techniques, random, low-level user-introduced contamination is unavoidable in work such as this. Each time a tube is uncapped, a pipette tip is used, or a reagent is opened, an opportunity for stochastic contamination arises. The depth of sequencing in the present study has allowed for such contamination to be revealed.

Salter et al. have demonstrated that reagent contamination adversely affects not only PCR-dependant methods, such as 16S rRNA gene sequencing, but also PCR-independent methods, such as shotgun metagenomics [74]. Our results reveal a low level of contamination in our workflow from the DNA extraction process. We were also unable to control for contamination introduced during library preparation and sequencing. Similarly, contamination from aspects of PCR-independent workflows, including DNA extraction and sequencing, may introduce low levels of contamination. These would be expected to vary between labs and reagents, and therefore should be identified and controlled for to limit inter-experiment variability in low-biomass microbiome studies.
Future studies should aim to test this method for eliminating contamination in a range of PCR reagents, as the efficiency and effectiveness of this treatment may vary based on the chemical properties of the specific PCR master mix.

**Summary**

In conclusion, we have demonstrated that, at least in our DNA extraction kit and PCR reagents, the major source of contamination in the microbiome workflow is the master mix (the “mixome”), rather than the extraction kit (“the kitome”). This contamination can be eliminated quickly and efficiently by treating master mixes with a dsDNase enzyme, allowing meaningful data to be produced in low biomass microbiome studies. The source of this contamination could be the master mix buffer, the *Taq* polymerase, dNTPs, MgCl$_2$ or the primers. It is important to note that manufacturers of these products do not guarantee absence of contaminating bacteria or DNA. If these products are produced non-aseptically, remnant bacterial DNA could cause significant problems in downstream universal bacterial screening applications. This type of contamination is especially problematic for studies of low-titre microbial communities, such as blood and meconium [72, 73, 252]. Limiting reagent-based contamination using dsDNase treatment could have a significant impact on the sensitivity and accuracy of microbiome studies performed on low biomass samples.
TABLE 4.2: 16S rRNA gene sequence reads produced from negative extraction controls (EC1-4) and negative PCR controls (NT1-2) that were amplified with or without dsDNase-treated master mix.

<table>
<thead>
<tr>
<th>RDP classification</th>
<th>dsDNase-treated master mix</th>
<th>Untreated master mix</th>
</tr>
</thead>
<tbody>
<tr>
<td>Genus</td>
<td>EC1</td>
<td>EC2</td>
</tr>
<tr>
<td>Escherichia/Shigella</td>
<td>33</td>
<td>63</td>
</tr>
<tr>
<td>Pelomonas</td>
<td>1</td>
<td>1788</td>
</tr>
<tr>
<td>Methylobacterium</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Sphingomonas</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Unknown Neisseriaceae</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Methylobacterium</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>Bradyrhizobium</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Novosphingobium</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Achromobacter</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Corynebacterium</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Enterobacter</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Ralstonia</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Parasegetibacter</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Rhodococcus</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Unknown Alphaproteobacteria</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Propionibacterium</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Micrococcus</td>
<td>0</td>
<td>315</td>
</tr>
<tr>
<td>Gemmobacter</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Cupriavidus</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bifidobacterium</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Chapter 5

The not-so-sterile womb: Evidence that the human fetus is exposed to bacteria prior to birth

This chapter addresses Aim 2 of this thesis - to characterise the bacterial composition of meconium and amniotic fluid at term in normal pregnancy. Having optimised DNA extraction methods for use with meconium and having validated a new method for minimising reagent-based contamination in the previous two chapters, here I have set out to perform the most rigorous study of the fetal microbiome published to date. I have compared fetal gut and amniotic fluid bacterial DNA profiles with meconium short chain fatty acid levels, amniotic fluid cytokine levels, and maternal health data. This analysis has contributed to Aim 3 of this thesis – to examine the relationship between maternal health, fetal bacterial colonisation, placental inflammation and fetal immune development. The data presented in this chapter support the in-utero colonisation theory, by demonstrating that the fetus is exposed to bacteria or bacterial DNA in uncomplicated pregnancies.

This chapter is currently in press in Frontiers in Microbiology as: Stinson LF, Boyce MC, Payne MS, and Keelan JA. 2018. The not-so-sterile womb: Evidence that the human fetus is exposed to bacteria prior to birth. Frontiers in Microbiology.
Abstract

The human microbiome includes trillions of bacteria, many of which play a vital role in host physiology. Numerous studies have now detected bacterial DNA in first-pass meconium and amniotic fluid samples, suggesting that the human microbiome may commence in-utero. However, these data have remained contentious due to underlying contamination issues. Here, we have used a previously described method for reducing contamination in microbiome workflows to determine if there is a fetal bacterial microbiome beyond the level of background contamination. We recruited 50 women undergoing elective caesarean section deliveries with no evidence of intra-uterine infection and collected first-pass meconium and amniotic fluid samples. Full-length 16S rRNA gene sequencing was performed using PacBio SMRT cell technology, to allow high resolution profiling of the fetal gut and amniotic fluid bacterial microbiomes. Levels of inflammatory cytokines were measured in amniotic fluid, and levels of immunomodulatory short chain fatty acids (SCFAs) were quantified in meconium. All meconium samples and most amniotic fluid samples (36/43) contained bacterial DNA. The meconium microbiome was dominated by reads that mapped to Pelomonas puraquae. Aside from this species, the meconium microbiome was remarkably heterogeneous between patients. The amniotic fluid microbiome was more diverse and contained mainly reads that mapped to typical skin commensals, including Propionibacterium acnes and Staphylococcus spp. All meconium samples contained acetate and propionate, at ratios similar to those previously reported in infants. P. puraquae reads were inversely correlated with meconium propionate levels. Amniotic fluid cytokine levels were associated with the amniotic fluid microbiome. Our results demonstrate that bacterial DNA and SCFAs are present in-utero, and have the potential to influence the developing fetal immune system.

Background

It has long been assumed that the human fetus is sterile. Over the last decade new data have emerged to challenge this dogma. However, these ideas are still controversial and there is disagreement amongst perinatal and microbiological researchers as to whether the human microbiome is seeded prior to birth. While numerous studies have reported the detection of bacterial [11, 13, 14, 24-26, 55-59, 62, 63, 257, 258], archaeal [13], fungal [13], and viral [62] DNA in meconium and amniotic fluid, the interpretation of
these data is contentious due to underlying contamination issues. It is now well accepted that laboratory reagents, including nucleic acid extraction kits and PCR master mix reagents, harbour lows levels of bacterial DNA [74]. While this contamination is not an issue for studies of highly colonised samples such as adult faeces, it becomes an important issue when working with low biomass samples such as meconium and amniotic fluid. Until recently, this contamination was largely dealt with by sequencing blank extraction and PCR controls and including these in the analysis. However, we have recently reported a method for removing contamination from PCR master mixes (the "mixome") which, combined with a carefully performed magnetic bead-based extraction, minimises contamination in our microbiome workflow to the point that it is barely detectable [259]. This allows more sensitive and accurate profiling of low biomass samples. Previous studies of the fetal microbiome are plagued with numerous other methodological issues. In particular, most of these studies use short amplicon sequencing, which is frequently unable to give accurate genus or species level identification [260]. Several of these studies also perform their analyses using QIIME v1, a pipeline that has been shown to give inaccurate results, with one study suggesting that 56-88% of genus names given by QIIME 1 are false positives [261]. Finally, several of these studies have not taken steps to account for the high levels of PCR inhibitors in meconium, which has led some investigators to conclude that meconium is sterile.

While studies in this field are complicated by methodological problems, those which have at the very least used DNA extraction and amplification controls in their sequencing analyses provide compelling evidence that the intra-amniotic space is unlikely to be sterile in healthy human pregnancies. In fact, if maternal-fetal transmission of microbes does not occur in human pregnancy, we would be one of the few animals studied in which this phenomenon did not occur [79].

Prenatal seeding of the human microbiome would be expected to have significant physiological implications for the developing fetus. The human microbiota produce important metabolites, including short chain fatty acids (SCFAs), which are absorbed systemically and which modify immune function and development [262]. The presence of bacteria in the intra-amniotic space might trigger immune activation or sensitisation,
resulting in the production of inflammatory mediators and other immune modulators. Maternal-fetal transmission of microbes during gestation would likely have a significant impact on the fetal immune system, gut, and brain. Additionally, this early scaffolding of the fetal microbiome could influence postnatal colonisation events. Thus, it is essential to conclusively determine whether the fetus harbours a microbiome, or whether the observed fetal microbiome is merely the result of contamination and other methodological errors.

The aims of this study were to characterise the bacterial DNA profiles of the fetal gut and amniotic fluid using full-length 16S rRNA gene sequencing and a previously described protocol for minimising reagent-based contamination [259], taking steps to eliminate environmental contamination as far as possible. Additionally, we profiled levels of inflammatory cytokines in the amniotic fluid and SCFAs in meconium to explore any potential metabolic-immunological-microbiological interactions.

**Methods**

**Patient recruitment and sample collection**

Patients giving birth by elective caesarean section between 34 and 42 weeks gestation (n = 50) at King Edward Memorial Hospital, Subiaco, Western Australia, were invited to participate in this study with the approval of the Human Research Ethics Committee of the Western Australian Department of Health’s Women and Newborns Health Service (2015212EW). Inclusion criteria were: singleton pregnancies, elective caesarean section deliveries, and a gestational age of ≥34+0 weeks. Indications for caesarean section delivery are presented in **Table 1**. Exclusion criteria were: onset of labour, antibiotic or antimycotic use throughout the pregnancy, antenatal steroid administration, vaginal progesterone administration, fetal genetic abnormalities, and recreational drug abuse. Participants answered a detailed questionnaire regarding their health, diet and lifestyle during their current pregnancy. Clinical data were also collected, including gestational age, maternal age, parity, infant sex, indications for caesarean section, previous obstetric history, and length of stay in neonatal intensive care unit (if any). Maternal and fetal characteristics are reported in **Table 2**. Five deliveries occurred at late preterm (34-37 weeks gestation). Amniotic fluid samples (n = 43) were collected aseptically during
caesarean section deliveries. Approximately 10 mL of fluid was drawn into a syringe via a 10 cm cannula fully inserted into the incision site immediately following amniotomy, then transferred to sterile tubes and centrifuged at 40,000 X g for 6 minutes at 4°C to pellet cells. Pellets were resuspended in 200 µL nuclease-free water (Integrated DNA Technologies) and both pellets and supernatants were frozen at -80°C until processing.

First-pass meconium samples (n = 50) were collected within the first 24 h of delivery (mean: 8.26 h; range: 0.67 – 18.55 h). In cases where neonates were fed prior to producing their first meconium, they had received colostrum only. Nappies were removed by gloved midwives and placed in sterile transport bags. Samples were then taken from the nappies in a laminar flow cabinet using aseptic techniques. A central portion of meconium was drawn into a sterile single use syringe to avoid contamination from external portions which may have contacted the neonate’s skin or nappy. Meconium samples were divided into 200 ± 3 mg aliquots and stored at -80°C until extraction.

**TABLE 5.1:** Indication for caesarean delivery in this cohort (n = 50). Multiple indications for caesarean delivery were common.

<table>
<thead>
<tr>
<th>Indication for caesarean section</th>
<th>Number of cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous caesarean section</td>
<td>35</td>
</tr>
<tr>
<td>Maternal diabetes</td>
<td>12</td>
</tr>
<tr>
<td>Elevated maternal BMI</td>
<td>5</td>
</tr>
<tr>
<td>Intra-uterine growth restriction</td>
<td>3</td>
</tr>
<tr>
<td>Tubal ligation</td>
<td>3</td>
</tr>
<tr>
<td>Tokophobia or previous traumatic vaginal delivery</td>
<td>3</td>
</tr>
<tr>
<td>Advance maternal age</td>
<td>3</td>
</tr>
<tr>
<td>Macrosomic fetus</td>
<td>2</td>
</tr>
<tr>
<td>Breech position</td>
<td>2</td>
</tr>
<tr>
<td>Placenta previa</td>
<td>2</td>
</tr>
<tr>
<td>Maternal congenital heart defect</td>
<td>1</td>
</tr>
<tr>
<td>Placenta accreta</td>
<td>1</td>
</tr>
<tr>
<td>Previous myomectomy</td>
<td>1</td>
</tr>
<tr>
<td>Colitis</td>
<td>1</td>
</tr>
</tbody>
</table>
**TABLE 5.2:** Maternal, fetal, and pregnancy characteristics for this cohort (n = 50). Values are reported as mean (range) or n (percent).

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean (range) or n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age (y)</td>
<td>32.7 (20-44)</td>
</tr>
<tr>
<td>Gravidity (n)</td>
<td>3.0 (1-11)</td>
</tr>
<tr>
<td>Parity (n)</td>
<td>1.6 (0-7)</td>
</tr>
<tr>
<td>GA at birth (week&lt;sup&gt;day&lt;/sup&gt;)</td>
<td>38&lt;sup&gt;±2&lt;/sup&gt; (34&lt;sup&gt;0&lt;/sup&gt;-42&lt;sup&gt;0&lt;/sup&gt;)</td>
</tr>
<tr>
<td>Pre-pregnancy BMI</td>
<td>29.4 (18.0-52.1)</td>
</tr>
<tr>
<td>Smoking in the three months prior to pregnancy</td>
<td>4 (8%)</td>
</tr>
<tr>
<td>Smoking during pregnancy</td>
<td>3 (6%)</td>
</tr>
<tr>
<td>T2 diabetes</td>
<td>6 (12%)</td>
</tr>
<tr>
<td>GDM</td>
<td>13 (26%)</td>
</tr>
<tr>
<td>Maternal history of asthma</td>
<td>11 (22%)</td>
</tr>
<tr>
<td>Maternal history of eczema/dermatitis</td>
<td>6 (12%)</td>
</tr>
<tr>
<td>Maternal history of allergies</td>
<td>7 (14%)</td>
</tr>
<tr>
<td>Male fetus</td>
<td>26 (52%)</td>
</tr>
</tbody>
</table>

*Short chain fatty acid profiling of meconium*

Meconium aliquots (200 mg) from 47 participants (insufficient sample volume in three cases) were transferred to 2 mL Eppendorf tubes. An acidified (0.05 M) 10 % aqueous methanol extraction solution spiked with 2-ethyl butyric acid as the internal standard (0.84 mM) was added and the tube was vortexed for 1 minute. The tube was left to stand for 60 minutes at 4°C and then vortexed again for 1 minute. The samples were centrifuged at 400 X g for 20 minutes at 4°C. Supernatant (1 mL) was transferred to a GC vial and analysed by GC-MS within 48 h. A blank control was prepared in the same way. Separation of SCFAs was achieved on a Thermo Fisher Scientific GC-MS (ISQ) using a Thermo Fisher Scientific TG-Wax column (30 m x 0.25 mm x 0.25 µm). The injector temperature was set at 230°C in splitless mode and an injection volume of 1 µL. The column temperature was initially 50°C for 0.5 minutes and then ramped to 110°C over 0.6 minutes. The temperature was then increased to 135°C at 4°C/minute and then ramped to 230°C at 20°C/minute and held for 5 minutes. The MS transfer line was maintained at 230°C and the ion source at 250°C. The MS was operated in electron impact mode and the mass range scanned was 50-250 amu. Quantification was based on a 7-point calibration curve. Acetic acid standards were prepared in the range 0.4-20 mM, butyric and propionic acid in the range 0.2-10 mM, isobutyric acid in the range
0.05-2 mM, isovaleric and valeric acid in the range 0.07-3.6 mM, and isohectanoic and hexanoic acid in the range 0.06-3 mM.

**Cytokine analysis of amniotic fluid**

Levels of interleukin 6 (IL-6), interleukin 10 (IL-10), C-X-C motif chemokine 10 (CXCL10), and granulocyte-colony stimulating factor (G-CSF) were measured in amniotic fluid supernatants by multiplex assay (Merck Millipore) on a MAGPIX instrument (Luminex Corp) as per the manufacturer’s instructions. The limits of detection for IL-6, IL-10, CXCL10, and G-CSF were 0.9, 1.1, 8.6, and 1.8 pg/mL, respectively.

**DNA extraction**

DNA was extracted from meconium (n = 43) and amniotic fluid pellets (n = 43) using a QIAGEN MagAttract DNA/RNA Microbiome kit (QIAGEN Pty Ltd, Hilden, Germany) on the King Fisher DUO platform (Thermo Fisher Scientific) as per manufacturer’s instructions for DNA, with the exception of the bead-beating step, which was performed in 2 mL tubes on a Precellys 24 bead beater at 6,500 RPM for 45 seconds. Seven out of the 50 collected meconium samples could not be extracted due to adverse reactions with an optional lysis component of the protocol (addition of phenol:chloroform:isoamyl alcohol, 25:24:1 to Solution MBL). This optional step was omitted from all other extractions. One blank extraction control was included in each batch of extractions (11 samples and one control per batch).

**Amplification and barcoding**

PCR was performed on DNA extracted from meconium and amniotic fluid samples to amplify the full 16S rRNA gene for sequencing. The primers 27F (5’-AGRGTYYGATYMTGCTCAG-3’) and 1492R (5’-RYTGACTTTGTTACGACTT-3’) were used (previously described here [263]), with the universal sequences UNITAG-F (gcgtgctagactagctgctgctgctg) and UNITAG-R (tggatctgtgcagcatgctgctgctg) incorporated at the 5’ end of these. An amine block (5’NH₄-C₆) was then incorporated at
the 5’ ends of each tagged primer to ensure that unbarcoded amplicons carried over from the first round of PCR did not undergo DNA ligation with SMRTbell adaptors during library preparation. A set of three forward and 15 reverse barcoded primers consisting of a unique barcode sequence (Pacific Biosciences (PacBio)) incorporated with either the UNITAG-F or -R sequence were designed to generate PacBio sequencing-ready amplicons using an asymmetric barcoding strategy (similar to that of [264]). All primers were synthesised and HPLC-purified by Integrated DNA Technologies.

To obtain barcoded 16S rRNA gene amplicons, amplification was carried out in two steps. The first PCR was carried out in 50 µL reactions containing 7.5 µL of template or water (negative template control), 1X AccuStart II ToughMix (Quantabio), 0.3 µM each of the forward and reverse primers, 1.25 µL each of the ArcticZymes dsDNase and dithiothreitol (DTT) (ArcticZymes), and 13.5 µL of water. PCR mastermix solutions were decontaminated before use using dsDNase as previously described [259]. The PCR amplification program consisted of an initial heating step at 94°C for 3 minutes; 40 cycles of 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 2 minutes; and a final extension step of 72°C for 7 minutes. PCR reactions were performed in a Veriti Thermal Cycler (Applied Biosystems). PCR products were visualised on a QIAxcel automated electrophoresis system using a DNA high resolution gel cartridge (run parameters OM500) to confirm the presence and size of amplicons.

Primary PCR products were purified using Agencourt AMPure XP magnetic beads (Beckman Coulter), normalised to 1 ng/µL, then used as template in a secondary, nested PCR, in order to generate asymmetrically barcoded amplicons. Secondary PCR was carried out in 25 µL reactions containing 2 µL of template or water (negative template control), 1X AccuStart II ToughMix (Quantabio), 0.3 µM each of the forward and reverse primers, and 3 µL of water. The PCR amplification program consisted of an initial heating step at 94°C for 3 minutes; 10 cycles of 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 2 minutes; and a final extension step of 72°C for 7 minutes.
**PacBio Sequencing**

Barcoded 16S rRNA gene amplicons obtained from the secondary PCR were pooled in equimolar concentrations based on QIAxcel quantitation of the target band (24 samples per pool). Pools were then concentrated using Agencourt AMPure XP magnetic beads and eluted in 50 µL volumes. Pools were next visualised in a 1.2% agarose gel using SYBR Safe DNA stain (Invitrogen) and bands of the appropriate size were excised using a sterile disposable scalpel (new scalpel for each band). Excised bands were purified using a QIAquick Gel Extraction kit (Qiagen) as per manufacturer’s instructions. 500 ng of purified DNA (amplicon) was used for library preparation of each pool (four pools in total) at the Queensland University of Technology Genomics Research Centre. Here, SMRTbell adapters were ligated onto barcoded PCR products, and the libraries were sequenced on a PacBio Sequel system (version 5.1.0) on a single SMRT cell (per pool).

**Sequence data analysis**

PacBio raw reads were processed using the SMRT Link Analysis software (version 6.0) to obtain demultiplexed circular consensus sequence (CCS) reads with a minimum of three full passes and 99.5% sequence accuracy. Sequence data were processed using the software package GHAP v2.1. GHAP is an in-house amplicon processing pipeline developed by Paul Greenfield (CSIRO, Australia) [253] built around tools from USEarch [229] and The Ribosomal Database Project (RDP) [254], combined with locally-written tools for generating OTU tables. BIOM files [265] generated by GHAP were analysed using MicrobiomeAnalyst – a web-based tool for statistical and visual analysis of microbiome data [231]. Reads were initially denoised using GHAP with a minimum number of three reads in a minimum of one sample required to retain an OTU. Minimum read filtering was applied in MicrobiomeAnalyst for alpha and beta diversity calculations, but was increased to a minimum of 10% prevalence with a count of 3 for differential abundance analysis. Alpha diversity was assessed as number of observed OTUs and Chao1. Beta diversity was visualised using PCoA plots and assessed using PERMANOVA. Differential abundance was calculated univariately. All P-values were calculated using student’s t-test unless stated otherwise.
Results & discussion

Amniotic fluid and first-pass meconium samples contain bacterial DNA

First-pass meconium samples were taken as a proxy for fetal gut contents in-utero. Previous studies have reported the detection of bacterial DNA in meconium samples; however, these data have been disputed due to the possibility of reagent-derived contamination. Here, we have sequenced the full-length 16S rRNA gene in first-pass meconium samples using a protocol that reduced reagent-based contamination to nearly undetectable levels [259]. Using these methods, we found all meconium samples to contain bacterial DNA. We recovered an average of 9,081 reads per meconium sample, which mapped to an average of 6.7 OTUs per sample (Supplementary table 5.1a). For each polymerase read, we achieved an average of 21 sequencing passes with an average read length of 1555 bp. Almost all meconium samples contained high numbers of reads that mapped to Pelomonas puraquae with a 99.5% sequence homology. Five samples did not follow this pattern (Figure 5.1). One of these was dominated by reads that mapped very poorly to an unknown Rhodospirillaceae, with only 77% sequence homology. Another harboured a relatively high level of diversity and contained mainly reads that mapped to Staphylococcus spp. (S. haemolyticus and S. epidermis, both skin commensals [266]), Streptococcus spp. (including S. infantis, an oral and nasopharyngeal commensal [267] and S. pseudopneumoniae ), Gemella taiwanensis (a newly described species isolated from human blood [268]), and Rothia mucilaginosa (an oral and upper respiratory tract commensal [269]). The third was dominated by reads from Lactobacillus iners (a vaginal commensal [270]). The fourth contained mainly reads from Pantoea agglomerans (commonly found in plants, and animal and human faeces [271-273]), and the fifth was dominated by reads that mapped to Cupriavidus gilardii (previously associated with sepsis [274, 275]).
We sequenced an average of 2,196 reads per amniotic fluid sample, which mapped to an average of 7.6 OTUs per sample (Supplementary table 5.1b). For each polymerase read, we achieved an average of 24 sequencing passes, with an average read length of 1532 bp. There were seven samples that produced a low number of reads (<50). Overall, amniotic fluid samples contained a low-abundance and low-diversity bacterial microbiome (Figure 5.2). Unlike meconium, only twelve amniotic fluid samples contained reads that mapped to *P. puraquae*. Other dominant reads included *Propionibacterium acnes*, a human skin commensal, which was also present in twelve samples. Six samples were dominated by reads that mapped to *Staphylococcus spp.* (*S. haemolyticus* and *S. lugdunensis*, both skin commensals). Three samples were dominated by reads that mapped to *Ralstonia pickettii*, a bacterium associated with water and fresh produce, as well as human pathogenic infections and nosocomial infections. One sample was dominated by reads that mapped to *Streptococcus anginosus*, a human commensal that exists in multiple body sites, including the mouth, gut, and vagina. Another was dominated by reads that belonged to *Peptoniphilus grossensis* – a newly discovered bacterium isolated from the faeces of a morbidly obese woman. Two of these species, *P. acnes* and *S. lugdunensis*, have previously been cultured from amniotic fluid samples from uncomplicated pregnancies [11].

We note that some of the taxa identified here are not biologically plausible human microbiome candidates. In particular, reads mapping to *Thermothrix azorensis* and *Thermus scotoductus*, thermophilic bacteria that are found in extreme environments, were found in very low numbers in two amniotic fluid samples (<100 reads) and in higher numbers in five meconium samples (122–1335 reads). Both of these sequences mapped to these taxa very well (>99% homology); however, in all cases the OTUs mapped just as well to “uncultured bacteria” as they did to these species. To further investigate these sequences we manually trimmed them to either side of the V4 region and classified them using basic local alignment search tool (BLAST) [276]. The taxonomic assignments for these OTUs remained as *T. azorensis*, *T. scotoductus*, and “uncharacterised bacteria” using this strategy. These sequences may, therefore, be misclassified, and instead represent currently uncharacterised bacteria which share a high level of sequence homology across the 16S rRNA gene.
FIGURE 5.1: Abundance (number of reads) of bacterial species detected in meconium samples.
FIGURE 5.2: Abundance (number of reads) of bacterial species detected in amniotic fluid samples.
Colonisation or contamination?

Almost all meconium samples were strongly dominated by reads that mapped to \textit{P. puraquae}. One explanation for the appearance of \textit{P. puraquae} in these samples is contamination in our workflow. The ubiquity of \textit{P. puraquae} in our meconium samples might suggest that its appearance is due to a shared source of contamination. Members of the \textit{Pelomonas} genera have previously been identified as contaminants in microbiome reagents [72, 74] and ultrapure water [255, 256]. Indeed, although \textit{P. puraquae} was completely absent in our PCR controls, it did appear in 1/5 of our meconium extraction controls and 2/5 of our amniotic fluid extraction controls (Table 5.3), so we cannot exclude the possibility of contamination. We have, therefore, presented Figures 5.1 and 5.2 with the dominant \textit{P. puraquae} OTUs removed in Supplementary Figures 5.1 and 5.2. However, \textit{P. puraquae} was absent in more than half of our amniotic fluid samples (31/43 samples), and in four of our meconium samples. Thus, its appearance in our extraction controls might be due to sporadic contamination from the samples themselves, as the extraction controls were placed in the centre of each extraction plate. The four neonates that lacked \textit{P. puraquae} DNA in their meconium also lacked \textit{P. puraquae} DNA in their amniotic fluid. This adds biological plausibility to the theory that the reads mapping to \textit{P. puraquae} truly were generated from bacterial DNA found in the fetal gut. Another possible source of contamination for these meconium samples is the neonate’s nappy. We did not take swabs of the nappies worn by these infants, but attempted to control for this type of contamination by sampling internalised portions of meconium only. If the \textit{P. puraquae} DNA detected in these meconium samples originated from the nappy, we would not expect to see it in any amniotic fluid samples, which we do. Therefore, we conclude that \textit{P. puraquae}, the primary read recovered from these meconium samples, was unlikely to be an external contaminant.

The majority of reads recovered from our amniotic fluid samples mapped to typical skin commensals, including \textit{P. acnes} and \textit{Staphylococcus spp}. It might be argued that these species are contaminants introduced during sampling or processing. However, it is highly unlikely that skin bacteria could have contaminated these samples during the sampling process. Maternal skin had been excised prior to amniotomy, and samples were drawn up into a newly opened, sterile syringe directly from within the amniotic sac. While we did not take blank controls during sampling, it is unlikely that these
samples were exposed to airborne contamination in the operating theatre, as the amount of time that they were exposed to air was negligible. Any user- or reagent-introduced contamination during sample processing would be expected to appear in the extraction and PCR controls that were taken. Apart from a single read that mapped to *P. acnes*, no other skin commensals appeared in these controls (Table 5.3). Therefore, we conclude that the amniotic fluid contains DNA from human skin commensals that are not representative of contaminants.

Of course, the presence of bacterial DNA in the intra-amniotic space does not prove “colonisation”. These bacteria may be transiently transferred from mother to fetus, where they might be unsuited to survive. However, exposure to bacteria and their by-products, even if they do not constitute a true microbiome in the ecological sense, might influence fetal immune development. Alternatively, cell-free DNA from these bacteria may be present in amniotic fluid and meconium samples and thus contribute to the microbial DNA signature. In the present study, however, cell-free DNA is unlikely to be a major contributing factor to the observed bacterial DNA profiles, as DNA is not readily pelleted by centrifugation unless this is done in liquid such as ethanol or isopropanol. Clarification of this question would require confirmation using bacterial culture methods, propidium monoazide, RNA in-situ probes, or cDNA-based sequencing. While not all previous studies have been able to culture live bacteria from meconium and amniotic fluid samples [46-48, 277], many have [11, 43-45, 58, 78]. However, it should be noted that in an environment as under-characterised as this, it would be nearly impossible to provide the growth requirements for all bacteria present.
TABLE 5.3: OTUs detected in blank controls taken during DNA extraction from meconium samples (Meconium extractions EC 1-5) and amniotic fluid samples (AF extractions EC 1-5). Percent homology to sequences reported on the RDP database is included. All blank PCR controls were completely negative and are therefore not included in this table.

<table>
<thead>
<tr>
<th>Taxonomic assignment</th>
<th>% match</th>
<th>Meconium extractions</th>
<th>AF extractions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>EC 1</td>
<td>EC 2</td>
</tr>
<tr>
<td>Pelomonas puraquae</td>
<td>99.5</td>
<td>8</td>
<td>2</td>
</tr>
<tr>
<td>Cloacibacterium rupense</td>
<td>95.8-98.5</td>
<td>3320</td>
<td></td>
</tr>
<tr>
<td>Cloacibacterium normanense</td>
<td>96.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown labedella</td>
<td>92.4-94.9</td>
<td>310</td>
<td></td>
</tr>
<tr>
<td>Pseudoclavibacter soli</td>
<td>93.3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cupriavidus gilardii</td>
<td>99.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ralstonia insidiosa</td>
<td>99.7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propionibacterium acnes</td>
<td>99.9</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown rhizobiales</td>
<td>91.9-93.4</td>
<td>28</td>
<td></td>
</tr>
</tbody>
</table>
First-pass meconium contains SCFAs

All meconium samples contained detectable levels of the SCFAs acetic acid and propionic acid, but not butyric acid, iso-butyric acid, valeric acid, iso-valeric acid, hexanoic acid, or iso-hexanoic acid (Figure 5.3). These samples contained an average of 29.35 mmol/g of acetic acid and an average of 4.37 mmol/g of propionic acid. Previous studies have found a similar ratio of acetic acid and propionic acid in infant stool collected on days 2, 3, 5 and 6 of life [278-280]. These studies were also able to detect very low levels of butyric acid. Given that our own samples were collected on the first day of life, butyric acid-producing bacteria may not colonise the gut until after birth. Alternatively, butyrate may be present in these samples, but below the limit of detection in this experiment (0.05 mM). Our data support previous observations that the SCFA profiles of infant stool differs from that of adult stool, which typically contains equal levels of propionic acid and butyric acid [280, 281].

FIGURE 5.3: Short chain fatty acid levels (mmol/g) in first-pass meconium samples (n = 47).
Is the presence of *P. puraquae* in these samples biologically plausible?

The most abundant sequence recovered from our meconium samples mapped to *P. puraquae*. Despite our best efforts, this species may represent contamination in our workflow. Here, we consider whether this result is biologically plausible. Importantly, these results differ considerably from those previously reported in the literature. Existing studies have reported that the meconium microbiome is dominated by members of the Enterobacteriaceae family (particularly *Escherichia* spp. and *Enterobacter* spp.), *Propionibacterium* spp., and lactic acid-producing bacteria (including *Lactobacillus* spp., *Leuconostoc* spp., *Enterococcus* spp., and *Lactococcus* spp.) [11, 13, 24-26, 55, 56]. *Pelomonas* spp. have never before been reported to be a major part of the meconium microbiome. Additionally, most previous studies report a higher level of within-sample and between-sample diversity. It is unusual to have samples that are so strongly dominated by reads from a single species. One explanation for these discrepancies is the fact that this is the first study of the meconium microbiome to be performed in Australia. Geographic variations in the human microbiome are well established, and Australia, being a relatively isolated country, may be unique in this regard [282-285]. In fact, *Pelomonas* sp. has previously been identified as a part of the Australian drinking water microbiome [286].

The apparently high level of diversity seen in previous studies of the meconium microbiome might also be due to methodological errors. In particular, several previous studies have used QIIME v1 for OTU clustering and taxonomic assignment. QIIME v1 has recently been shown to overestimate the number of OTUs in a given sample. In fact, up to 88% of predicted genus names given by QIIME v1 may be false positives [261]. Contamination may also account for some of the bacteria that have previously been reported as part of the meconium microbiome. *Escherichia* spp. are often described as a major part of the meconium microbiome; however, we have previously shown that *Escherichia* sp. DNA is a major contaminant of PCR reagents, possibly due to remnant DNA from the *E coli* used to produce *Taq* DNA polymerase [259]. Here, we have decontaminated our PCR reagents using a dsDNase treatment, which we have previously shown to dramatically reduce reagent-based contamination by 99% [259]. Further, in studies in which the samples used were not first-pass meconium samples, or in which samples were collected days after birth, sample contamination from the *ex utero* environment may occur. Neonates are rapidly colonised by bacteria from
breastmilk and the environment following birth, so it is essential to take only early, first-pass meconium samples to use as proxies of the fetal gut at birth [49]. In the present study, all meconium samples used were passed in the first day of life, an average of 8.26 h following birth. However, previous studies have included samples passed multiple days after birth, casting doubt on their ability to represent the fetal gut contents at birth. Additionally, previous studies of the meconium microbiome have largely relied on short amplicon sequencing, while we have used full-length 16S rRNA gene sequencing. Short amplicon sequencing lacks the taxonomic sensitivity and specificity of full-length amplicon sequencing, and can result in the generation of false positives [260]. Short amplicon sequencing can be more efficient at producing amplicons in some samples due to DNA fragmentation. On the other hand, while full-length primers generate amplicons that provide superior taxonomic data, they generally have reduced coverage compared to some short amplicon primers. The primers used here cover 81.2% of full-length 16S rRNA gene sequences for the domain Bacteria in the SILVA database, given 1 mismatch within 5 bases of the 3’ end. Based purely on primer coverage, we cannot dismiss the possibility that the results presented here may be an underestimation of the full bacterial microbiome of these samples. Additionally, short PCR amplicons are typically generated more easily than long PCR amplicons.

We are not the first to identify Pelomonas spp. as a coloniser of the naïve human gut. A recent study performed in Crohn’s disease patients tracked microbial assemblage and recolonisation following ileocolonic resection and colonic cleansing. These authors found that patients who had recurrence of the disease following surgery had distinctly different recolonisation processes than those who were in disease remission. For those patients who experienced disease recurrence, bacterial recolonisation followed one of five trajectories. One of these trajectories was affiliated with OTUs belonging to the Comamonas and Pelomonas genera [287]. These findings support the notion that Pelomonas spp. may be able to colonise naïve human gut environments, and also support other research showing that Pelomonas is a dominant genera in the mucosa-attached microbiota of patients with severe Crohn’s disease [288]. Pelomonas spp. have also been identified in human blood [289], in the oral microbiome [290], and in endometrial samples [291, 292]. None of the aforementioned studies, however, controlled for reagent-based contamination, or were able to achieve species level resolution (all short amplicon sequencing studies), so it is difficult to make conclusions
on the occurrence of *P. puraquae* in human tissue. Nevertheless, *Pelomonas spp.* DNA has been recovered from non-pregnant endometrial samples, potentially indicating that the uterus is the source of the bacterial DNA that we recovered in meconium and amniotic fluid samples.

**Presence and abundance of *P. puraquae* in meconium is inversely correlated with meconium propionate levels**

Given the ubiquity of *P. puraquae* in these samples, we divided the meconium samples based on whether *P. puraquae* was absent, present but not dominant, or dominant. For this purpose we defined *P. puraquae* dominance as meaning that the number of reads assigned to *P. puraquae* was greater than the sum of all other reads. This allowed us to compare characteristics based on the presence and abundance of *P. puraquae*. Meconium samples that were dominated by reads belonging to *P. puraquae* had significantly lower concentrations of propionic acid than those that were not dominated by *P. puraquae* (*P* = 0.016). Similarly, *P. puraquae* presence in meconium samples, regardless of whether or not it was dominant, was associated with significantly lower levels of propionic acid in these samples also (*P* = 0.00055). *P. puraquae* is able to grow on acetate [293]; however, its ability to utilise or produce propionate has not been described.

**Detection of *P. acnes* and *P. puraquae* in amniotic fluid is associated with altered amniotic fluid cytokine profiles and microbial richness**

*P. acnes* is a ubiquitous human skin commensal. The full genome of *P. acnes* has been sequenced, revealing its potential to produce immunogenic proteins [294]. Thus, its presence in the amniotic fluid could suggest that fetuses are exposed to bacterial-derived immune stimulation prior to birth. Interestingly, the twelve mothers who harboured *P. acnes* DNA in their amniotic fluid were found to have lower levels of amniotic fluid CXCL10 and IL10 (*P* = 0.058 and 0.025) (Figure 5.4a). However, *P. acnes* presence within the amniotic fluid was also associated with a significant increase in amniotic fluid microbial richness (number of observed OTUs, *P* = 0.029). The observed reduction in cytokine levels in *P. acnes* colonised amniotic fluid may, therefore, be due to tolerance to a range of bacteria.
*P. puraquae* DNA was detected in twelve amniotic fluid samples and was the dominant read in eight of these. G-CSF was elevated in amniotic fluid samples that contained *P. puraquae* DNA (*P* = 0.018) (Figure 5.4b). Amniotic fluid samples in which *P. puraquae* DNA was present, but not dominant, had a significantly higher level of alpha diversity (number of observed OTUs) than samples in which *P. puraquae* was the dominant read (*P* = 0.0021). Given that *P. puraquae* may be an unavoidable contaminant in our workflow, we removed this OTU from our dataset and re-analysed the relationship between amniotic fluid cytokine levels and bacterial load, alpha diversity, and beta diversity. We were unable to detect any such relationship.

**FIGURE 5.4:** A: Amniotic fluid cytokine levels (pg/mL) from pregnancies with *P. acnes* DNA present (pink) or absent (blue) in the amniotic fluid. B: Amniotic fluid cytokine levels (pg/mL) from pregnancies with *P. puraquae* DNA present (pink) or absent (blue) in the amniotic fluid. * *P* = 0.058, ** *P* = 0.025, *** *P* = 0.018.
Meconium might be seeded from the amniotic fluid microbiome

Given that the fetus swallows amniotic fluid throughout the second and third trimesters, the amniotic fluid and meconium microbiomes may be expected to share a large portion of their microbiota. Indeed, previous studies have shown that the meconium microbiome shares more features with the amniotic fluid microbiome than with the maternal faeces, placenta, colostrum, or infant faeces [11]. Here, we found that 32.7% of detected species were shared between amniotic fluid and meconium samples, while 28.6% were found in the amniotic fluid only and 38.8% were found in the meconium only (Figure 5.5a). However, the 32.7% of shared species accounted for 93.6% of total reads, while only 2.4% of reads were found in the amniotic fluid only and only 4.0% of reads were found in meconium only (Figure 5.5b). Given that we cannot eliminate the possibility that *P. puraquae* is a contaminant in our workflow, we re-analysed our OTU table without this species. Without *P. puraquae* 72.6% of reads are shared between the two sample types. Although 16S rRNA gene-based microbiome studies are only semi-quantitative at best, these data indicate that the majority of bacterial DNA found in the fetal gut is also found in the amniotic fluid.

While the majority of reads recovered mapped to species that were shared between the two sample types, there were distinct differences in the relative abundances of these reads (Figure 5.6). Univariate analysis revealed that *P. puraquae* was significantly more abundant in meconium samples (Mann Whitney U test, \( P = 4.39 \times 10^{-10} \)), while *P. acnes* was significantly more abundant in amniotic fluid samples (Mann Whitney U test, \( P = 0.0078 \)) (Figure 5.7).
FIGURE 5.5: A: Venn diagram showing species that were unique to meconium or amniotic fluid samples, or shared between them. B: Venn diagram showing number of reads that were unique to meconium or amniotic fluid samples, or shared between them.
FIGURE 5.6: Relative abundance of bacterial genera recovered from meconium and amniotic fluid (AF) samples.
Interestingly, *P. puraquae* was strongly dominant in meconium samples, but was only present in twelve amniotic fluid samples and only dominant in eight of these. Importantly, subjects that lacked *P. puraquae* in their meconium samples also lacked it in their amniotic fluid samples. One hypothesis that may explain these observations is that *P. puraquae* may have a competitive advantage over other bacteria in the fetal gut. Meconium is a very unusual substance, made up of bile acids and salts, pancreatic secretions, epithelial cells, and the residue of swallowed amniotic fluid. *P. puraquae* might be well suited to this environment, while other bacteria might be outcompeted or unable to survive in meconium. Similarly, amniotic fluid samples were largely
dominated by skin commensals such as *P. acnes* and *Staphylococcus* spp., which may be suited to such a niche, as amniotic fluid contains high levels of sloughed fetal skin cells [295]. Thus, these data may demonstrate bacterial niche development *in-utero*.

**Meconium and amniotic fluid microbiomes do not vary with maternal health parameters**

Aside from the dominant *P. puraquae* sequences, the bacterial DNA profiles of meconium and amniotic fluid in this cohort were remarkably heterogeneous between individuals. We collected a range of maternal health and lifestyle data from the women in our cohort; however, neither beta diversity nor OTU presence/abundance varied with any of these parameters (including maternal pre-pregnancy BMI, parity, gravidity, maternal diabetes, maternal asthma or atopy, maternal smoking, maternal ethnicity, and gestational age at delivery). This was true whether or not *P. puraquae* was removed from the analysis. These results stand in contrast to previous studies which have found that both the amniotic fluid [63] and the meconium [56] microbiomes vary with maternal diabetes status. Lack of power may be a factor here: our cohort included six mothers with type II diabetes mellitus (T2DM) and thirteen with gestational diabetes (GDM). Although no significant differences in the microbiota were observed in these women, we did find that neonates born from mothers with T2DM had significantly lower levels of propionic acid in their meconium compared to those born from mothers with normal pancreatic function (*P* = 0.0048) or from mothers with GDM (*P* = 0.00328). These differences may be driven by the maternal microbiome. The gut microbiota have been implicated in the pathophysiology of diabetes, with the production of SCFAs suggested as a mechanistic pathway linking microbes and T2DM [296]. Thus, mothers with T2DM would be expected to harbour altered gut microbiomes and altered circulating SCFA profiles. These SCFAs might be transferred across the placenta to the fetus.

It has been theorised that fetal microbiome establishment might vary in relation to maternal atopic disease [297]. Gosalbes *et al*. reported shifts in the meconium microbiota that were related to maternal eczema, but not to asthma or rhinitis [24]. In the present study, we did not find any difference in alpha diversity, beta diversity or
community membership based on maternal asthma, allergies, hay fever, or eczema. However, small cohorts such as the one used here lack the power to detect changes in “normal” microbiota [298].

**Summary**

Here we have provided the first full-length 16S rRNA gene survey of meconium and amniotic fluid. Our data suggest that the fetus is exposed to bacterial DNA and metabolites prior to birth. Additional studies that utilise pre-extraction sample treatment with membrane-impermeable dyes such as propidium monoazide to prevent PCR amplification of DNA contained within compromised cell walls will help to shed light on whether this DNA originates from viable or dead bacteria and as such whether it constitutes a true microbiome.
SUPPLEMENTARY FIGURE 5.1: Abundance (number of reads) of bacterial species detected in meconium samples presented with the two dominant P. puraqua OTUs removed.
SUPPLEMENTARY FIGURE 5.2: Abundance (number of reads) of bacterial species detected in amniotic fluid samples presented with the two dominant *P. puraqua* OTUs removed.
Chapter 6
Detection of bacterial DNA in mid-gestation amniotic fluid samples

This chapter addresses Aim 5 of this thesis – to determine whether amniotic fluid contains bacterial DNA in mid-pregnancy and, if so, to characterise the taxonomic profile of these samples. To date, only six studies have investigated the amniotic fluid microbiome using next generation sequencing, and only three of these included mid-gestation amniocentesis samples. There is, therefore, a need for a properly controlled investigation into the mid-gestation amniotic fluid microbiome. Here we utilise the dsDNase method described in Chapter 4 to decontaminate PCR reagents before screening 1198 mid-gestation amniocentesis samples by 16S rRNA gene qPCR. A selection of those that screened positive were then subjected to full 16S rRNA gene sequencing. The data presented in this chapter indicate that bacterial DNA begins to enter the amniotic fluid as early as mid-gestation in some pregnancies with no evidence of clinical infection.
Abstract

Controversy exists as to the timing and route of seeding of the human microbiome in early life. To date there has been very little investigation into the amniotic fluid microbiome using next generation sequencing technology. In particular, if microbes are able to access the amniotic space during pregnancy, there is a critical gap in our knowledge of the timing of such an event. Here, 1198 mid-gestation amniotic fluid samples were screened for bacterial DNA using qPCR in a cohort of Swedish women undergoing amniocentesis for fetal genetic investigation in mid-gestation (14-20 weeks). The majority of samples were devoid of detectable bacteria; however, approximately a fifth of the cohort (19.9%) were 16S rRNA gene positive in duplicate screening. We used PacBio SMRT cell technology to sequence the full-length 16S rRNA gene in a subset of 18 of these samples from women who went on to deliver at term. These amniotic fluid samples contained low-abundance and low-diversity bacterial DNA profiles that were dominated by reads that mapped to Saccharibacteria (TM7 oral clone), Acidovorax temperans, Tepidimonas taiwanensis, and Pelomonas puraquae. These findings provide evidence that, in a minor proportion of pregnancies, DNA from non-pathogenic bacteria is present in the amniotic fluid far earlier than previously reported.

Background

There is an increasing body of evidence to suggest that the establishment of the human microbiome begins in-utero [11, 210]. Numerous studies have investigated bacterial colonisation of the amniotic fluid in complicated and healthy pregnancies using culture-dependant and culture-independent techniques. These studies have reported rates of amniotic fluid bacterial colonisation that range from 0-100% [11, 14, 62, 63, 108, 133, 139, 277, 299-305]. While pathogenic colonisation of the amniotic fluid has been heavily studied in the context of preterm birth and fetal infection [133, 139], little is known about the nature, origins, and significance of any endemic amniotic fluid bacteria in normal pregnancies. In particular, the timing and stability of this colonisation remains largely unexplored. Such information is important, as bacterial populations in the amniotic fluid are likely to influence fetal skin colonisation and fetal gut colonisation via swallowing of the amniotic fluid which begins at the end of the first trimester.
To date, six next generation sequencing studies of amniotic fluid from uncomplicated pregnancies have been published. Collado et al. sequenced the 16S rRNA gene in amniotic fluid samples from 15 mothers delivering at term by elective caesarean section [11]. While their study yielded interesting results, including the detection of viable bacteria in the amniotic fluid by culture, it is complicated by the fact that appropriate extraction and PCR controls were not used. Furthermore, these authors used QIIME v1 for OTU clustering and taxonomic assignment. QIIME v1 has previously been shown to give inaccurate taxonomic assignments and to vastly overestimate the number of OTUs in a sample [261]. More recently, Lim et al. sequenced the 16S rRNA gene in amniotic fluid samples from 24 healthy women delivering by caesarean section [62]. These authors also performed virome sequencing on these samples. Contrary to the title of their paper, “Amniotic fluid from healthy term pregnancies does not harbour a detectable microbial community”, their data showed that bacteria and a diverse range of bacteriophage were detectable in all pregnancies. However, the authors instead interpreted their results as showing that amniotic fluid is indistinguishable from controls, despite their supplementary data clearly showing that numerous OTUs were present in amniotic fluid at levels far exceeding that of the controls. This study did not attempt to eliminate reagent-derived contamination using laboratory-based techniques, such as those discussed in Chapter 4 of this thesis, or bioinformatic techniques, such as those described by Davis et al. and de Goffau et al. [240, 306]. This makes it difficult to make conclusions on the true level of bacterial colonisation in their samples. Wang et al. sequenced the 16S rRNA gene in amniotic fluid samples from 23 women with gestational diabetes mellitus (GDM) and 41 healthy controls delivering by caesarean section [63]. They reported that the amniotic fluid bacterial microbiome is clearly distinct in pregnancies that are complicated by GDM compared to those that are not. Unfortunately, these authors also used QIIME v1 for OTU clustering and taxonomic assignment and did not utilise blank extraction or PCR controls. Urushiyama et al. sequenced the 16S rRNA gene in amniotic fluid samples from 79 caesarean section deliveries and 18 mid-gestation amniocentesis samples [14]. These authors sought to identify predictive microbiome profiles for chorioamnionitis. They found that healthy pregnant women harboured a rich amniotic fluid bacterial microbiome at delivery and that chorioamnionitis severity was inversely correlated with alpha diversity. A strength of this study was its use of negative controls for extraction and library preparation; however, these authors did not attempt to eliminate reagent-derived contamination.
Kayem et al. screened 1043 amniotic fluid samples taken for genetic testing at 16-20 weeks gestation [300]. These authors reported that 31 of these were 16S rRNA gene positive by qPCR, 29 of which were also positive upon qPCR screening for *Ureaplasma spp.* and *Mycoplasma spp.*. Bacterial DNA was successfully sequenced in five samples that were positive by qPCR. These samples were made up of *Sphingomonas sp.*, *Acinetobacter sp.*, *Bacillus sp.*, and *Streptococcus sp.*. No negative extraction or PCR controls were used. Similarly, Zhu et al. screened 64 amniotic fluid samples taken for genetic screening at 17-20 weeks gestation by 16S rRNA gene qPCR [302]. These authors found all 64 samples to be positive; however, they could not cultivate bacteria from any of these samples. Sequencing of these samples revealed bacterial profiles made up of Enterobacteriaceae, *Propionibacterium sp.*, Bacillales, *Anoxybacillus sp.*, Caulobacteraceae, Methylobacteriaceae, *Methylobacterium sp.*, *Phyllobacterium sp.*, *Sphingomonas sp.*, *Achromobacter sp.*, Comamonadaceae, Xanthomonadaceae, and *Deinococcus sp.*. Unfortunately, these authors did not utilise negative extraction or PCR controls and used QIIME v1 to process their sequence data.

Apart from being inadequately controlled, these studies have all relied on short amplicons and different variable regions for their sequencing. While such an approach is somewhat useful for broad community profiling without detailed taxonomic information, it is not able to give accurate species level assignment [260]. Additionally, the selection of the region of the 16S rRNA gene to be sequenced can introduce bias into the results [307, 308]. The best way to overcome these issues is by sequencing the full 16S rRNA gene, as was previously done prior to next generation sequencing technologies using molecular cloning, albeit with vastly reduced sequencing depth. Until recently, such an approach had been rarely used, as it has been prohibitively expensive and fraught with random sequencing errors. The only option available (prior to Oxford Nanopore) was the use of PacBio single molecule real-time (SMRT) sequencing, a technology more commonly used for long-read sequencing as part of whole genome studies. However, PacBio has recently developed the next iteration of SMRT DNA sequencing, known as circular consensus (CCS) reads. This technology has resulted in a huge increase in sequencing accuracy for amplicons up to 2 kB in size, with estimates of >99% accuracy being reported, comparable to Illumina and Ion Torrent short read technologies. In addition, samples can be multiplexed with unique
barcode identifiers resulting in reduced costs comparable to those of short read sequencing technologies [264].

The existence of an amniotic fluid microbiome is contentious; however, increasing evidence suggests that bacterial DNA is commonly found in amniotic fluid at birth [11, 14, 63]. It is not known when in gestation this colonisation might be first established. Apart from the previously described studies of Urushiyama et al., Kayem et al., and Zhu et al. [14, 300, 302], studies of the mid-gestation amniotic fluid microbiome to date have relied on bacterial culture and targeted real-time PCR, and have primarily focused on identifying pathogenic bacteria involved in the aetiology of preterm birth [108, 301, 305, 309]. High quality data on the full microbial ecology of mid-gestation amniotic fluid is lacking, making it difficult to speculate on the extent to which bacterial DNA is found in amniotic fluid at this time point.

There is currently a need for a robust, well-controlled analysis of the bacterial content of amniotic fluid in normal pregnancies, particularly one which employs full-length 16S rRNA gene sequencing, coupled with appropriate measures to reduce reagent contamination. In this study we characterised the microbial profile of amniotic fluid collected from mid-gestation amniocentesis from pregnancies with no clinical evidence of infection that went on to deliver at term.

**Methods**

**Sample collection**

Mid-gestation amniotic fluid samples (14-20 weeks, n = 1198) were collected by amniocentesis with the approval of the Central Ethics Review Board at the University of Gothenburg, Sweden (Dnr Ö 639-03). Written informed consent was obtained from all participants. Indications for amniocentesis in this cohort were advanced maternal age, anxiety, abnormal first-trimester combined screening or family history of chromosomal abnormalities or genetic diseases. The exclusion criteria were maternal age of <18 years, multiple pregnancy, positive HIV or hepatitis B test and known or suspected fetal
malformation. The medical and demographic details of this cohort are presented in Table 6.1.

3 mL of amniotic fluid was collected for research purposes (in addition to that collected for diagnostic use) during mid-gestation genetic transabdominal amniocentesis and immediately stored at 4-8°C. The samples were centrifuged at 12,000 X g for 20 minutes at 4°C to pellet the cells. Pellets were resuspended in 1 mL sterile 1X PBS and frozen at -80°C until processing.

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Mean (range) or n (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Maternal age (y)</td>
<td>36.4 (20-47)</td>
</tr>
<tr>
<td>Gravidity</td>
<td>2 (0-10)</td>
</tr>
<tr>
<td>Parity</td>
<td>1.2 (0-7)</td>
</tr>
<tr>
<td>GA at birth (week^day)</td>
<td>39^4 (22^1-43^10)</td>
</tr>
<tr>
<td>Pre-pregnancy BMI</td>
<td>24.5 (15.8-47.4)</td>
</tr>
<tr>
<td>Smoking during pregnancy</td>
<td>62 (5.2%)</td>
</tr>
<tr>
<td>T2 diabetes</td>
<td>4 (0.33%)</td>
</tr>
<tr>
<td>GDM</td>
<td>15 (1.3%)</td>
</tr>
<tr>
<td>Male fetus</td>
<td>592 (49.4%)</td>
</tr>
</tbody>
</table>

**DNA extraction**

Amniotic fluid suspensions were centrifuged at 40,000 X g for 5 minutes at 4°C to pellet cells. Cell pellets were resuspended in 353 µL Buffer MBL/RNase A solution (QIAGEN) and DNA was extracted using a QIAGEN MagAttract Microbial DNA kit on the KingFisher Flex platform as per manufacturer’s instructions. One blank extraction control was included in the centre of each 96 deep-well plate.
qPCR Screening

All mid-gestation samples were screened for bacterial DNA using a custom Taqman qPCR protocol, adapted from Yang et al. with a new forward primer to improve coverage [225]. Briefly, the V6 region of the 16S rRNA gene was amplified in 20 µL reactions containing 6 µL of template or water (negative template control), 1X TaqMan Fast Advanced Master Mix (Applied Biosystems), 0.9 µM each of the forward (917F 5′-GAATTGACGGGGRCCCGC -3′) and reverse (1033R 5′-TGCGGGACTTAACCCAACAA -3′) primers, 0.25 µM of probe (5′-FAM-CACGAGCTGACGACARCCATGCA-TAMRA’), and 0.95 µL of water. Master mix solutions were treated with a PCR Decontamination Kit (ArcticZymes®), which consisted of a double-stranded DNase (dsDNase) and DTT (which helps to inactivate the dsDNase). Briefly, master mix solutions (including primers and probes) were treated with 0.5 µl of dsDNase and 0.5 µl of DTT per 20 µl reaction, then incubated at 37˚C for 20 minutes (dsDNase activation), followed by incubation at 60˚C for 20 minutes (dsDNase inactivation). After this, decontaminated master mix solutions were cooled on ice for two minutes to prevent heat-related changes in volume. This treatment eliminated amplification of background microbial DNA, preventing negative template controls from reading as positive, and allowing positive/negative calls to be made confidently. Each qPCR was run with three negative template controls and two human DNA controls (1 ng and 10 ng), as human DNA is a known confounder in low biomass 16S rRNA gene studies due to 16S rRNA gene primers/probes binding to human gDNA with varying affinity [310]. For a sample to be called positive, it had to have a Ct value at least 2 cycles less than the human DNA controls. In this way we were able to account for mis-priming with human DNA. Samples were screened in duplicate to confirm the findings, and only samples that produced positive results for both replicates were considered positive.

We selected 18 mid-gestatation samples that tested 16S rRNA gene positive for subsequent sequencing. These were samples from pregnancies that were not complicated by infection or inflammation and that went on to deliver at term. These samples were selected as demographically-matched controls for a case-control study of
the amniotic fluid microbiome in preterm birth, reported on elsewhere (manuscript in preparation).

Amplification and barcoding

Amniotic fluid samples from 18 pregnancies that went on to deliver at term and were positive for bacterial DNA by qPCR screen were selected for full-length 16S rRNA gene sequencing. The primers 27F (5'-gcagtcgaacctgtagctgactcagtcacAGRGTTYGATYMTGGCTCAG-3') and 1492R (5' tggatcacttgtgcaagcatcacatcgtagRGYTACCTTGTTACGACTT-3') were used, with the universal UNITAG sequences (in lower case) and a 5' (NH₄-C₆) block attached to each primer. A set of three barcoded UNITAG-F and 15 barcoded UNITAG-R primers were designed to generate PacBio sequencing-ready amplicons, using an asymmetric barcoding strategy. All primers were synthesised and HPLC-purified by Integrated DNA Technologies.

To obtain barcoded 16S rRNA gene amplicons, amplification was carried out in two steps. The first PCR was carried out in 50 µL reactions containing 7.5 µL of template or water (negative template control), 1X AccuStart II ToughMix (Quantabio), 0.3 µM each of the forward and reverse primers, 1.25 µL each of the ArcticZymes dsDNase and DTT, and 13.5 µL of water. The PCR amplification program consisted of an initial heating step at 94°C for 3 minutes; 40 cycles of 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 2 minutes; and a final extension step of 72°C for 7 minutes. PCR reactions were performed in a Veriti Thermal Cycler (Applied Biosystems). PCR products were visualised on a QIAXcel automated electrophoresis system using a DNA high resolution gel cartridge (run parameters OM500) to confirm the presence and size of amplicons.

Primary PCR products were purified using Agencourt AMPure XP magnetic beads (Beckman Coulter), normalised to 1 ng/µL, then used as template in a secondary, nested PCR, in order to generate asymmetrically barcoded amplicons. Secondary PCR was carried out in 25 µL reactions containing 2 µL of template or water (negative template control), 1X AccuStart II ToughMix, 0.3 µM each of the forward and reverse primers,
and 3 µL of water. The PCR amplification program consisted of an initial heating step at 94°C for 3 minutes; 10 cycles of 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 2 minutes; and a final extension step of 72°C for 7 minutes.

**PacBio Sequencing**

Barcoded 16S rRNA gene amplicons obtained from the secondary PCR were pooled in equimolar concentrations based on QIAxcel quantitation of the target band. The pool was then concentrated using Agencourt AMPure XP magnetic beads and eluted in 50 µL of 1X TE. The pool was next visualised in a 1.2% agarose gel using SYBR Safe DNA stain (Invitrogen) and bands of the appropriate size were excised using a sterile disposable scalpel. Excised bands were purified using a QIAquick Gel Extraction kit as per manufacturer’s instructions. 500 ng of purified DNA was used for library preparation at the Queensland University of Technology Genomics Research Centre. Here, SMRTbell adapters were ligated onto barcoded PCR products, and the library was sequenced on a PacBio Sequel system on a single SMRT cell.

**Sequence data analysis**

PacBio raw reads were processed using the SMRT Link Analysis software version 6.0 to obtain demultiplexed circular consensus sequence (CCS) reads with a minimum of three full passes and 99.9% sequence accuracy. Sequence data were processed using the software package GHAP v2.1. GHAP is an in-house amplicon processing pipeline developed by Paul Greenfield (CSIRO, Australia) [253] built around tools from USEarch [229] and The Ribosomal Database Project (RDP) [254], combined with locally-written tools for generating OTU tables. BIOM files [265] generated by GHAP were analysed using MicrobiomeAnalyst – a web-based tool for statistical and visual analysis of microbiome data [231]. Reads were initially denoised using GHAP with a minimum number of three reads in a minimum of two samples required to retain an OTU.

**Results & discussion**
Bacterial DNA is present in amniotic fluid in mid-pregnancy

Of the 1198 mid-gestation amniotic fluid samples screened by qPCR, 238 (19.9%) contained detectable levels of bacterial DNA. The average cycle threshold for the 16S rRNA gene positive samples was 34.14 (range: 28.51 – 43.88). Using the methods described in Chapter 4 of this thesis, we were able to produce completely negative extraction and PCR controls in our qPCR screen, allowing us to make positive/negative calls with a higher level of sensitivity and accuracy. We are not the first to screen amniotic fluid samples from mid-pregnancy using universal 16S rRNA gene primers. Two previous such studies have reported that 0% and 3% of samples were positive for bacterial DNA [300, 301]. Neither of these studies, however, reported any details of negative controls used nor the method used for distinguishing positive samples from background noise. Without the use of decontaminated reagents, negative controls frequently appear at cycle thresholds similar to those of amniotic fluid in 16S rRNA gene qPCR screening tests. It is, therefore, possible that these studies lacked the level of sensitivity necessary to make conclusions on such low biomass samples. Further, the primers used by Kayem et al. and Rowlands et al. covered only 21.5% and 40.0% of bacterial sequences in the SILVA database, respectively, while our own set covered 66.5% of these sequences.

Interestingly, our detection of bacterial DNA in the amniotic fluid was not associated with infection-mediated preterm birth. Only 14 of these 238 positive samples (6.1%) were from women who went on to deliver preterm (<37 weeks gestation), while 6.9% of the samples that tested negative for bacterial DNA were from women who went on to deliver preterm. There was also no difference in the average gestational age at sampling for those that tested positive for bacterial DNA compared to those that tested negative (15±6 weeks in both groups). Rates of other pregnancy complications (including intra-uterine growth restriction, pre-eclampsia, and GDM) were even lower than the rate of preterm birth in this cohort. Therefore, it was not possible to make statistically sound comparisons on the rates of these conditions in women who screened positive and those who screened negative for bacterial DNA. These data indicate that in some pregnancies with uneventful deliveries, bacterial DNA is present in the amniotic fluid at 14-20
weeks gestation. However, in the large majority of cases, the amniotic fluid appears sterile at this gestational age.

The potential exposure of the fetus to bacteria at this early gestational age is of particular interest, as the fetal skin does not keratinise until approximately 20 weeks gestation [311]. Prior to this stage the fetal skin offers no resistance to movement of the amniotic fluid, which can be thought of as an extension of the fetal extra-cellular fluid. After fetal skin keratinisation occurs, the amniotic fluid changes in osmolality and can no longer equilibrate with the fetus. The effect that fetal skin keratinisation has on the amniotic fluid microbiome is unknown. The samples examined in the present study were taken prior to gestational week 20. The fetal skin would not have acted as an effective barrier between the amniotic fluid microbiome and the fetal extracellular fluid at this stage in development.

**Mid-gestation amniotic fluid harbours a low abundance and low diversity bacterial profile**

Overall, the amniotic fluid samples profiled here contained a low-abundance and low-diversity bacterial profile, as was expected based on previous studies [11, 14, 62, 63] (Figure 6.1). Of the 18 amniotic fluid samples sequenced, four returned insufficient reads (≤50). Each of these samples gave a clear positive signal on the qPCR screen, but failed to produce an amplicon with the full-length 16S rRNA gene primers used for sequencing, suggesting that the bacteria present in these samples were not covered by these primers. This raises the possibility that the full breadth of the amniotic fluid microbiome was not captured in this study. Of the remaining samples, we recovered an average of 6,292 reads per sample (range 1,284 – 9,239) and 10 OTUs per sample (range 1-32). Here, our negative PCR controls were completely negative, while our negative extraction controls were sporadically positive. Six OTUs were detected in six extraction controls: *Pelomonas puraquae, Casalitella massiliensis, Staphylococcus pasteuri, Bosea enae, Acinetobacter beijerinckii, and Rhodobacter blasticus*. However, in cases where one of these OTUs was detected in an extraction control, it was not detected in the samples that were extracted in that batch, suggesting that extraction-
based contamination was not a contributor to our results. Such a result may also indicate sporadic contamination, although this is less likely.

At the species level, amniotic fluid samples were dominated by reads that mapped to *Tepidimonas taiwanensis*, *Acidovorax temperans*, Saccharibacteria (TM7 oral clone), and *Pelomonas puraquae* (Figure 6.2). Reads mapping to *P. puraquae* did so with a high level of sequence homology (99.5%). *P. puraquae* is a relatively uncharacterised bacteria, first isolated in 2007 from haemodialysis water [293], and previously reported in meconium samples and full-gestation amniotic fluid samples taken during caesarean delivery (Chapter 5). It is a Gram-negative, motile rod with a single polar flagellum. Although *P. puraquae* is often thought of as a contaminant in microbiome research, it was present in only one of the eight extraction controls taken here and in neither of the two PCR controls. The extraction controls processed in the same batch as the samples that contained reads mapping to *P. puraquae* did not contain reads mapping to *P. puraquae* themselves. Although this species has not previously been identified in mid-gestation amniotic fluid samples, Urushiyama et al. reported finding reads that mapped to *Pelomonas saccharophila* in mid-gestation amniotic fluid samples using short read sequencing [14].

Six samples contained reads that mapped to Saccharibacteria (TM7 human oral clone), an extremely small coccus (200-300 nm) that until recently has been uncultivated, but is ubiquitous in numerous environments [312]. It is important to note that Saccharibacteria was not a particularly good match for this particular OTU, with only 92.6-93% sequence homology. This might indicate that these OTUs belong not to Saccharibacteria, but to a related, as-yet undescribed bacteria (for example, an uncultivated clade of TM7). Four samples contained reads that mapped to *A. temperans* with a reasonably good level of sequence homology (98.3%). *A. temperans* is a Gram negative rod with mega-pili on its surface, and has previously been isolated from various clinical and environmental samples, including urine, cervical swabs, and wastewater [313]. Interestingly, all recorded cases of clinical isolation of *A. temperans* have been in Europe (Sweden, France, and the UK) [313, 314] and most of these have been in Gotenberg, Sweden [313] – the very city from which these samples were taken.
This species has not previously been identified in amniotic fluid. Three samples contained reads that mapped to *T. taiwanensis* with 97.2% sequence homology. *T. taiwanensis* is a relatively uncharacterised bacteria that was first described in 2006 in hot springs [315]. Similarly to *P. puraquae*, it is a Gram-negative, motile rod with a single polar flagellum. Given that *Tepidimonas taiwanensis* is a thermophile found at moderately high temperatures (50-60°C) it is unlikely to be the correct identity for this particular OTU. Although the sequence matched reasonably well to this particular bacterium, a 2.8% difference in a 1.5 kb amplicon represents a 42 bp difference in the 16S rRNA gene. It is, therefore, possible that this OTU belongs to an as yet uncharacterised bacterium. This OTU was not detected in any of the negative controls.

Other less abundant OTUs identified here are in line with those previously identified in mid-gestation amniotic fluid samples by others. Three of our samples contained reads that mapped to *Propionibacterium acnes*, which was previously identified in mid-gestation amniotic fluid by Zhu et al. and Urushiyama et al. [14, 302]. We also identified reads that mapped to *Bacillus* sp. (which could not be differentiated between *B. anthracis*, *B. cereus*, *B. mycoides*, *B. pseudomycoides*, *B. thuringiensis*, and *B. weihenstephanensis*) and *Bacillus niabensis*. Reads that map to species of the *Bacillus* genus have been consistently been reported in previous studies of the mid-gestation amniotic fluid microbiome [14, 300, 302]. Other genera detected here that have also been detected using short sequence technology in mid-gestation amniotic fluid include *Arthrobacter*, *Streptococcus*, *Staphylococcus*, *Massilia*, and *Paenibacillus* [14, 300, 302].

Our findings suggest that in the majority of pregnancies, the amniotic cavity is devoid of bacteria in mid-gestation; however, in a minor proportion low-level exposure to a wide range of bacteria occurs long before term. It is important to note that these data reflect the presence of bacterial cell-associated DNA in amniotic fluid, but are not necessarily indicative of viable cells or true bacterial colonisation. The lack of any consistent pattern of microbial colonisation (within our study and between studies) might argue against the notion that this represents an early amniotic fluid microbiome,
although this interpretation is constrained by the limited sample number and methodological differences between studies.

FIGURE 6.1: Absolute abundance (number of reads) of OTUs in 18 selected PCR-positive mid-gestation amniotic fluid samples and the genera they map to (n = 18).
**Limitations and directions for future research**

Although a large number of samples were screened for bacterial DNA here, only 18 were used for 16S rRNA gene profiling. Given the high level of variation seen in bacterial composition in these samples, it is likely that we have not captured the true level of diversity in this cohort. Future studies should aim to include a higher number of samples for bacterial community profiling.

A previous study of the bacterial and viral microbiome of full-gestation amniotic fluid samples demonstrated the presence of a large range of bacteriophages in these samples [62]. This exciting finding might suggest that there is more to the putative amniotic fluid microbiome than bacteria alone. It would, therefore, be informative to survey all...
microbial components of amniotic fluid throughout gestation to establish timeframes of colonisation for each and the dynamics between them.

Summary

Here we have established that, while most mid-gestation amniotic fluid samples are sterile, bacterial DNA is present in amniotic fluid at this time point in approximately one in five pregnancies. Importantly, by eliminating reagent-based contamination, we are the first to provide evidence that this early detection of bacterial DNA in amniotic fluid is not due to contamination alone, as others have previously concluded [62]. The developmental significance of the presence or absence of bacterial DNA in mid-gestation amniotic fluid samples is still unclear.
Chapter 7

Characterisation of the bacterial microbiome in first-pass meconium using propidium monoazide (PMA) to exclude non-viable bacterial DNA

DNA-based microbiome studies performed using 16S rRNA gene sequencing are limited by their inability to discriminate between live bacterial cells, dead bacterial cells, and cell-free DNA. In this chapter we use propidium monoazide to exclude non-viable bacteria from microbiome analysis of first-pass meconium samples and thereby reveal that the majority of the purported fetal gut microbiome is from intact bacterial cells. This work demonstrates the importance of excluding non-viable bacteria when analysing the microbial community in low biomass samples such as meconium. This chapter contributes to Aim 4 of this thesis - to ascertain whether the apparent fetal gut microbiome represents living, metabolically active cells, or dead cells and bacterial products transferred from the mother.

Abstract

Numerous studies have reported bacterial DNA in first-pass meconium samples, suggesting that the human gut microbiome is seeded prior to birth. However, these studies have not been able to discriminate between DNA from living bacterial cells, DNA from dead bacterial cells, or cell-free DNA. Here we have used propidium monoazide (PMA) together with 16S rRNA gene sequencing to determine whether there are intact bacteria cells in the fetal gut. DNA was extracted from first-pass meconium (n = 5) and subjected to 16S rRNA gene sequencing with/without PMA treatment. All meconium samples, regardless of PMA treatment, contained detectable levels of bacterial DNA; however, treatment with PMA prior to DNA extraction decreased DNA yield by around 20%. PMA-treated meconium samples did not differ significantly from untreated samples in terms of observed number of OTUs (P = 0.945), although they did differ taxonomically, with around one quarter of OTUs identified in untreated samples only, suggesting they originated from cell-free/non-viable DNA. The mean Sørensen coefficient for treated vs. untreated samples was 0.527. Our findings suggest that the fetal gut is seeded with intact bacterial cells prior to birth. This is an important finding, as exposure to live bacteria during gestation might have a significant impact on the developing fetus.

Background

In recent years, the long-held theory that the womb is sterile has been met with an increasing body of evidence to the contrary. Numerous 16S rRNA gene next generation sequencing (NGS) studies have reported on a fetal gut microbiome [11, 13, 24, 53-59]. However, while these studies have detected bacterial DNA in the fetal gut, they have not been able to provide information on whether this DNA comes from viable organisms. This is a critical gap in the literature. Is there a true, living, metabolically active fetal microbiome that affects host physiology prior to birth, or are dead bacterial cells, bacterial products, and cell-free bacterial DNA transferred to the fetus via the maternal circulation?
Studies on the fetal gut microbiome rely on first-pass meconium samples as a proxy for the fetal gut contents. Meconium is a problematic substance to work with as it is difficult to solubilise, contains high levels of PCR inhibitors, and is difficult to culture bacteria from [252]. To date there have been two culture-based studies that were able to detect bacteria in meconium [58, 78]. Jimenez and colleagues collected meconium samples that were passed within the first two hours of life and before feeding had commenced (n = 21). The outer layer of the meconium was removed to avoid contamination from the perianal skin or nappy. They were able to culture between one and five bacterial species from each sample, with *Enterococcus faecalis*, *Staphylococcus epidermidis*, and *Escherichia coli* being the predominant species. Similarly, Moles et al. collected 14 first-pass meconium samples and were able to culture bacteria from 11 of them [58]. Five genera were isolated from these samples, including *Staphylococcus* sp., *Enterococcus* sp., *Streptococcus* sp., *Lactobacillus* sp., and *Klebsiella* sp. Currently, these are the only culture-based studies of the fetal gut microbiome. Culture-based studies of the placenta [11, 137, 316], cord blood [64], and amniotic fluid [11] support the notion that live bacteria can exist in the womb. However, culture-based studies can underestimate the diversity of a microbiome and can fail to detect fastidious bacteria [3]. It is, therefore, necessary to use molecular techniques to selectively characterise the viable bacterial microbiome. These may include the use of membrane-impermeable DNA intercalating dyes, DNases, or metatranscriptomics [77].

A popular method for discriminating between living and dead bacteria in microbiome research is the use of propidium monoazide (PMA) [317]. PMA is a DNA modifying dye that binds to dsDNA with a high affinity. Photolysis with visible light causes PMA to become covalently bound to dsDNA, preventing it’s amplification by PCR. PMA is cell membrane impermeable; thus, in a population of live and dead cells, only DNA from live cells with intact cell membranes will be amplified by subsequent PCR. This method has previously been used to selectively amplify DNA from intact bacterial cells in the adult human gut [318], the preterm infant gut [319], water [320], and even space dust [321].
Here we have sought to provide further evidence that the fetal gut harbours a viable, living microbiome in normal healthy pregnancies using PMA and NGS.

**Methods**

**Sample collection**

Five first-pass meconium samples were collected from healthy infants at King Edward Memorial Hospital, Western Australia with the approval of the Human Research Ethics Committee of the Western Australian Department of Health’s Women and Newborn Health Service (2015026EW). All samples were passed within 12 h of birth (mean 6.4 h) and processed within 1 h of being passed. Whole nappies were removed from the infants by gloved midwives, de-identified and placed in sterile transport bags. Samples were then taken from the nappies in a level two biosafety cabinet using aseptic techniques. Samples were stored at -20°C until extraction.

**PMA treatment**

Meconium samples (200 ± 3 mg) were dissolved in nuclease-free water (Integrated DNA Technologies, Inc.) at a 1:10 ratio (w:v). Two 200 µL aliquots of meconium slurry were taken from each sample, one of which was treated using PMA (PMAxx™, Biotium) immediately prior to DNA extraction, and one of which was untreated. PMA treatment was performed following the manufacturer’s protocol. Briefly, samples were topped up to 500 µL with water. 1.25 µL of 20 mM PMAxx™ was added to each sample to a final concentration of 50 µM. Samples were vortexed for 20 seconds then incubated in the dark at 37°C for 15 minutes with occasional vortexing. Samples were then exposed to light using a PMA-Lite LED Photolysis Device (Biotium) for 15 minutes, with vortexing every 5 minutes to ensure all parts of the sample were exposed to light. Cells were pelleted by centrifuging at 5,000 X g for 10 minutes and were resuspended in 200 µL of nuclease-free water.
DNA Extraction

DNA was extracted from treated (n = 5) and untreated (n = 5) meconium samples using a QIAGEN MagAttract Microbial DNA kit on the Kingfisher Duo platform and eluted into 100 µL of elution buffer. An extraction control consisting of 200 µL of nuclease-free water was processed with and without PMA treatment.

DNA quantitation

Total DNA yield was assessed using the Qubit® dsDNA HS Assay kit with a Qubit® 2.0 fluorometer. The limit of detection was 10 pg/µL.

16S rRNA gene sequencing

The V3-V4 region of the 16S rRNA gene was amplified using the primer pair 341F (5’-CCTACGGGNGGCGAGCAG-3’) and 785R (5’-GACTACHVGGGTATCTAATCC-3’), previously validated as providing optimal coverage of the domain Bacteria for a 400-1000 bp amplicon [227]. PCR was carried out in 50 µL reactions containing, 1X AccuStart II ToughMix (Quantabio), 0.3 µM each of the forward and reverse primers, 15 µL of template or water (negative template control) and water to bring the volume to 50 µL. PCR master mixes were treated with the Arcticzymes PCR Decontamination Kit prior to the addition of template as per manufacturer’s instructions to remove any contaminating microbial DNA prior to amplification. The PCR amplification program consisted of an initial heating step at 94°C for 3 minutes, followed by 40 cycles of 94°C for 30 seconds, 55°C for 30 seconds, and 72°C for 2 minutes, and a final extension step of 72°C for 7 minutes. PCR reactions were performed on an Applied Biosystems Veriti Thermal Cycler. PCR products were visualised on a QIAxcel automated electrophoresis system using a DNA high resolution gel cartridge (run parameters OM500) to confirm the presence and size of amplicons. PCR products were purified using NucleoMag® NGS Clean-up beads (Macherey-Nagel) following the manufacturer’s protocol and re-suspended in 20 µL of Low TE buffer (0.1 mM EDTA, 10 mM Tris-HCl, pH 8).
DNA library preparation was carried out using the Ion Plus Fragment Library Kit following the manufacturer’s protocol. Briefly, purified amplicons (100 ng) were blunt-ended using the End Repair Enzyme Mix. The Ion P1 Adaptor and Ion Xpress Barcodes were ligated to the amplicons while simultaneous nick repair was carried out. The adaptor-ligated libraries were purified using NucleoMag® NGS Clean-up beads, re-suspended in 20 µL of Low TE Buffer and amplified using the Platinum PCR SuperMix High Fidelity and Library Amplification Primer Mix. The thermal cycling conditions consisted of an initial denaturation at 95°C for 5 minutes, followed by 5 cycles of 95°C for 15 seconds, 58°C for 15 seconds and 70°C for 1 minute. The libraries were again purified using NucleoMag® NGS Clean-up beads and quantified using the Qubit dsDNA HS Assay Kit. Each library was adjusted to 100 pM in Low TE Buffer and combined in an equimolar ratio to ensure equal representation of each barcoded library in the sequencing reaction.

Automated template preparation using isothermal amplification technology and chip loading was performed using the Ion 520 & 530 ExT Kit on the Ion Chef System (Thermo Fisher Scientific). A 50 µL aliquot of the 100 pM pooled library was added to the Ion S5 ExT Reagents cartridge for templating onto Ion Sphere Particles (ISPs) and loading onto an Ion 530 Chip. The loaded Ion 530 Chip was sequenced for 1300 flows using the Ion S5 ExT Sequencing Kit on an Ion S5 Sequencer with Torrent Suite Software version 5.2.2 using Default Calibration (Thermo Fisher Scientific).

Data processing

Demultiplexed reads were quality and length filtered, clustered, and given taxonomic assignments using GHAP v2.1. GHAP is an in-house amplicon processing pipeline developed by Paul Greenfield (CSIRO, Australia) [253] built around tools from USEarch [229] and RDP [254], combined with locally-written tools for demultiplexing and generating OTU tables (de novo strategy). BIOM files [265] generated by GHAP were analysed using MicrobiomeAnalyst – a web-based tool for statistical and visual analysis of microbiome data [231]. Beta diversity analysis was performed group-wise using weighted UniFrac distances with similarity between PMA-treated and untreated
samples calculated using Permutational multivariate analysis of variance (PERMANOVA). Beta diversity analysis was also performed sample-wise using Sørensen’s coefficient. Alpha diversity measurements were based on number of observed OTUs and differences between PMA-treated and untreated samples were calculated using the student t-test.

**Results & discussion**

The overall yield of DNA from both the treated and untreated meconium samples was low, as expected based on previous reports [13, 252] (Table 7.1). Treatment with PMA prior to DNA extraction decreased the apparent yield of total DNA by an average of 20.5%. The attached dye molecules in PMA-conjugated nucleic acids can interfere with binding and fluorescence output of the quantification reagents (according to the manufacturer of these products, Thermo Fisher Scientific), resulting in a decrease in apparent yield that is relative to the decrease in unbound DNA. This might suggest that around 80% of the DNA present in meconium is from intact cells. It should be noted that the decrease in fluorescent output may also be due to differences in extraction efficiency of PMA-bound and unbound DNA. To more accurately quantify the bacterial DNA content of these samples qPCR could be used. However, qPCR quantification of DNA in meconium is highly inaccurate, as meconium contains high and variable levels of PCR inhibitors. Previous studies have reported a similar ratio of live to dead cells in adult stool using fluorescence-activated cell sorting and propidium iodide staining with flow cytometry, with dead cells making up 17% – 34% of the total sample [322, 323].
**TABLE 7.1:** Quantity of total DNA in each meconium sample with (+) or without (-) PMA treatment.

<table>
<thead>
<tr>
<th>Sample</th>
<th>PMA treatment</th>
<th>DNA quantity (ng/µL)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meconium 1</td>
<td>+</td>
<td>0.0623</td>
</tr>
<tr>
<td>Meconium 1</td>
<td>-</td>
<td>0.0640</td>
</tr>
<tr>
<td>Meconium 2</td>
<td>+</td>
<td>0.0500</td>
</tr>
<tr>
<td>Meconium 2</td>
<td>-</td>
<td>0.0616</td>
</tr>
<tr>
<td>Meconium 3</td>
<td>+</td>
<td>0.0830</td>
</tr>
<tr>
<td>Meconium 3</td>
<td>-</td>
<td>0.1090</td>
</tr>
<tr>
<td>Meconium 4</td>
<td>+</td>
<td>0.0420</td>
</tr>
<tr>
<td>Meconium 4</td>
<td>-</td>
<td>0.0592</td>
</tr>
<tr>
<td>Meconium 5</td>
<td>+</td>
<td>0.1600</td>
</tr>
<tr>
<td>Meconium 5</td>
<td>-</td>
<td>0.2220</td>
</tr>
<tr>
<td>Extraction control</td>
<td>+</td>
<td>Below limit of detection</td>
</tr>
<tr>
<td>Extraction control</td>
<td>-</td>
<td>Below limit of detection</td>
</tr>
</tbody>
</table>

Meconium samples varied in their community composition and the degree to which PMA treatment modified the resulting microbial profiles. In particular, the degree of similarity between treated and untreated samples varied from sample to sample, with Sørensen coefficients (a measure of β-diversity) ranging from 0.35 – 0.71 (mean = 0.53). There was no significant difference in the overall bacterial community composition of PMA-treated and untreated meconium samples (PERMANOVA R-squared: 0.10626; \( P < 0.25 \)); however, there were differences in individual OTU presence. PCOA analysis revealed that PMA treatment did not seem to affect sample 2, while it had an observable impact on the other samples (**Figure 7.1**).
FIGURE 7.1: Principle Coordinates Analysis based on weighted UniFrac distance. Individual meconium samples (n = 5) are labelled and colour coded based on whether they were PMA-treated prior to amplification and sequencing (pink) or untreated (blue).

Alpha diversity (number of observed OTUs) was decreased by PMA treatment in three samples (samples 1, 3 and 5) and increased after PMA treatment in two samples (samples 2 and 4) (Figure 7.2). The mean alpha diversity of the PMA-treated samples was not statistically different from that of the untreated samples (T-test $P = 0.945$).
Eight OTUs appeared only in untreated meconium samples (*Haemophilus* sp., *Dermacoccus* sp., *Streptomyces* sp., *Leptotrichia* sp., *Aquabacterium* sp., *Chryseobacterium* sp., *Porphyromonas* sp., and *Anaerococcus* sp.), suggesting that no viable cells of these genera exist in these samples (Figure 7.3). Interestingly, eight OTUs appeared only in PMA-treated samples (an unknown Microbacteriaceae, an unknown Actinomycetales, *Propionibacterium* sp., *Lacibacterium* sp., *Kocuria* sp., Rhodospirillaceae sp., *Curvibacter* sp., and *Amnibacterium* sp.). Given that amplicons were heavily diluted prior to sequencing, it might be that some low abundance OTUs were missed by chance. Alternatively, it is possible that high abundance, specific bacterial sequences were preferentially amplified in the untreated samples, and that a reduction in these following PMA-treatment allowed primers to bind to low abundance template that was previously unbound. This possibility is in line with an observation made by Rogers *et al.* that PMA treatment of sputum samples increased evenness and diversity, a phenomenon driven by increased detection of rare taxa [324]. Our findings confirm the presence of DNA from intact *Escherichia/Shigella* sp., *Streptococcus* sp., and *Lactobacillus* sp. cells in first-pass meconium, as reported by Jimenez *et al.* and Moles *et al.* using culture-dependant techniques [58, 78]. Additionally, we have
identified 24 OTUs that represent intact bacterial cells in meconium and 8 OTUs that represent dead bacteria or cell-free bacterial DNA.

The biological relevance of cell-free bacterial DNA in the fetal gut is likely to be quite different to that of metabolically active viable bacterial cells. TLRs which recognise ligands associated with bacterial cells (including lipopeptides, lipopolysaccharide, flagellin, and propellin) far outnumber those which recognise single or double stranded DNA (eg. TLR9). Therefore, a broader and more robust immunological response would be expected if the fetal gut is exposed to whole bacterial cells. This response will likely be modified depending on bacterial load, fetal gut barrier function, and the maturity of the fetal gut mucosal immune system. Exposure to live bacteria during gestation would be expected to have a significant impact on the developing fetus via the production of microbial metabolites and activation of innate immune sensing pathways, potentially priming the fetal immune system for life outside the womb. Our findings confirm that there are intact bacterial cells and bacterial DNA in first-pass meconium, thereby adding support to the theory that the fetal gut contains live bacteria and that the amniotic cavity is not a sterile environment.

Our findings reiterate the importance of taking into account cell viability in DNA-based microbiome studies, particularly those which interrogate environments that are believed to be largely sterile [325]. Microbiome studies using metagenomic techniques typically make conclusions about microbial ecology, physiology, pathology and function based on the assumption that the genomic profile is derived from viable microbiota. In many situations, this assumption is reasonable; however, in low biomass samples, particularly where evidence of sterility exists, the presence of cell-free DNA and dead bacteria can significantly confound the analysis and biological interpretation of microbiome data. In the present study, we identified eight OTUs that were present only in the non-PMA-treated samples. Were we to have examined these samples without viability testing, we may have made inaccurate physiological conclusions based on the presence of these bacteria. Our findings suggest that, in low biomass samples such as meconium, exclusion of DNA from non-viable bacteria should be performed to obtain accurate and biologically meaningful microbiome data.
FIGURE 7.3: The relative abundance (circle size) of the genera (y-axis) in each sample (x-axis) is plotted for PMA-treated (PMA) and untreated (UT) meconium samples.

<table>
<thead>
<tr>
<th>Genus</th>
<th>PMA 1</th>
<th>UT 1</th>
<th>PMA 2</th>
<th>UT 2</th>
<th>PMA 3</th>
<th>UT 3</th>
<th>PMA 4</th>
<th>UT 4</th>
<th>PMA 5</th>
<th>UT 5</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pseudomonas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Escherichia/Shigella</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Tepidimonas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propionibacterium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Staphyloccus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown Neisseriaceae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bradyrhizobium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Acinetobacter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dermacoccus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Micrococcus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Enhydrobacter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Corynebacterium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown Actinomycetales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Haemophilus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aquabacterium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Chryseobacterium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown Enterobacteriaceae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Lactobacillus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alloprevotibacter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown Actinomycetales</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kocuria</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Amnibacterium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unknown Microbacteriaceae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Methylobacterium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anaerococcus</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Leptotrichia</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Porphyromonas</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Streptomyces</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Curvibacter</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Propionibacterium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Locibacterium</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rhodospirillaceae</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
The use of PMA for viability testing in microbiome studies does, however, have its limitations. The effectiveness of PMA treatment varies with different bacterial species and in different sample types [326]. Viability assays such as the one used here therefore require careful optimisation and validation based on the target bacteria. However, in complex samples that contain multiple bacteria, such optimisation and validation is not feasible. For this reason, we used the default protocol for this study, which could potentially have under- or over-estimated the quantity of non-viable cells in these samples. Additionally, PMA-based viability assays are based on the assumption that all cells with intact membranes are viable without actually proving biological activity. Beyond PMA, metatranscriptomics may be used to test viability in microbiome studies. However, the use of metatranscriptomics in microbiome studies is limited by the difficulty of extracting high quality RNA and by the inherent instability of RNA. Degradation of RNA can result in underestimation of the viable microbial community in a sample, particularly in low biomass samples such as meconium where most bacteria are present in low quantities to begin with.

Viability testing prior to NGS is broadly applicable to other low biomass or supposedly sterile environments where it is important to determine the presence of intact bacterial cells. Future studies in this field should go beyond sequencing bacterial DNA and focus instead on sequencing DNA from intact bacterial cells by pre-treating samples with PMA or similar agents. Metabolomics should also be used to investigate possible interactions between the fetal gut microbiome and the developing fetal immune system.

**Summary**

Here we have provided evidence that the fetal gut contains DNA from intact bacterial cells. Our data support previously published culture-based investigations of the meconium microbiome. This study reinforces the importance of viability testing when analysing microbial communities in low biomass samples such as meconium.
Chapter 8

Placental and intra-amniotic inflammation are associated with altered fetal immune responses at birth

This chapter addresses Aim 3 of this thesis – to examine the relationship between maternal health, fetal microbial colonisation, placental inflammation and fetal immune development. Almost half of the placentae in this cohort showed histological evidence of inflammation. I have, therefore, investigated whether exposure to low-level placental inflammation was associated with altered immune responses in cord blood samples taken at birth. Further, I explored the extent to which bacterial DNA in the amniotic fluid and fetal gut might relate to the placental inflammation and fetal immune responses observed here.

Part of this chapter has been submitted for publication to Human Reproduction as: Stinson LF, Payne MS, and Keelan JA. 2019. Placental and intra-amniotic inflammation are associated with altered fetal immune responses at birth. Human Reproduction.
Abstract

High-grade placental inflammation is associated with preterm birth and poor neonatal outcomes. Recent reports suggest that low-grade placental inflammation is common in uncomplicated pregnancies. The relationship between placental inflammation and fetal immune programming is unknown. In this study we sought to identify any association between placental inflammation and fetal immune responses. Cord blood samples collected from late preterm and full-term caesarean section deliveries (n = 44) were exposed to various immune challenges (resiquimod, LPS, PGN, poly (I:C), cGAMP, and 5’ppp-dsRNA) and production of inflammatory mediators (G-CSF, IFN-γ, IL-1β, IL-6, IL-8, IL-10, and TNF-α) was measured by multiplex assay. Hospital histology reports were used to assess the extent of inflammation in the placenta. Almost half (47.7%) of placentae examined here showed histological evidence of inflammation. Resiquimod, LPS, and PGN elicited strong inflammatory responses in fetal cord blood, while poly (I:C), cGAMP, and 5’ppp-dsRNA elicited weaker responses. Fetuses with evidence of chorioamnionitis and fetal inflammatory reaction in their placentae had significantly increased immune responses to cGAMP and 5’ppp-dsRNA (ligands for STING and RIG-I, respectively) and significantly decreased immune responses to poly (I:C) (a TLR3 agonist). Interestingly, STING, RIG-I, and TLR3 are all involved in viral response pathways, suggesting that fetuses exposed to chorioamnionitis or fetal inflammatory reaction might respond differently to viruses postnatally. Our data suggest that low-level placental inflammation is associated with aberrant immune responses at birth.

Introduction

Placental inflammation is typically associated with infection, preterm birth, and poor neonatal outcomes [327]. Placental inflammation can be broadly categorised according to the site and extent of histopathological inflammation [328-333]. Placental inflammation includes deciduitis (DEC), subchorionitis, chorionitis, chorioamnionitis (CAM), chorionic vasculitis, umbilical vasculitis, funisitis, and villitis (including villitis of unknown aetiology; VUE). Different inflammatory lesions within the placenta are associated with different pathologies, aetiologies, and clinical outcomes. The three
major chronic inflammatory lesions of the placenta are CAM, VUE, and DEC [334]. CAM is a common cause of preterm birth, and is characterised by maternal neutrophil and CD8+ T-cell infiltration of the placental membranes [329]. Bacterial and fungal infections of the amniotic cavity are frequently identified in cases of CAM. VUE is characterised by infiltration of maternal lymphocytes into the chorionic villi, and can be associated with destruction of the villous architecture by macrophages [333]. VUE is less common in placentae delivered preterm. DEC is characterised by the presence of lymphocytes or plasma cells in the decidua basalis. Both infectious and sterile immune mechanisms have been implicated in this condition.

While placental inflammation has been extensively studied in the context of fetal and pregnancy pathologies, such as preterm birth and intra-uterine growth restriction, the prevalence and significance of inflammation in placentae from uncomplicated pregnancies has received less attention [335-338]. A recent study by Romero et al. revealed that placental inflammation is common in healthy, full-term pregnancies. These authors examined placentae from 944 uncomplicated full-term pregnancies for histological evidence of inflammation and found that 42.3% of these had acute inflammatory lesions, while 29.9% had chronic inflammatory lesions [339]. Interestingly, most of these cases were mild, with severe inflammation present in only 3.4% of cases. These data suggest that low-level placental inflammation is common in uncomplicated, full-term pregnancies. The consequence of exposure to low-level placental inflammation with respect to fetal immune development is unclear.

Fetal immune programming begins prior to birth, and in-utero events can alter postnatal immune responses [340, 341]. Intrauterine infection and CAM have been correlated with immune-related disorders, including an increased risk of asthma and atopic disease [342-344]. We therefore hypothesised that exposure to placental inflammation would alter neonatal immune responses to a range of immune stimuli.
Methods

Patient recruitment and ethics

Patients giving birth by elective caesarean section between 34 and 42 weeks gestation (n = 44) at King Edward Memorial Hospital, Subiaco, Western Australia, were invited to participate in this study. Cord blood and placental samples were taken at birth with the approval of the Human Research Ethics Committee of the Western Australian Department of Health’s Women and Newborns Health Service (2015212EW). Inclusion criteria were: singleton pregnancies, caesarean section deliveries, and a gestational age of ≥34⁰ weeks. Exclusion criteria were: onset of labour, antibiotic or antimycotic use throughout the pregnancy, antenatal steroid administration, vaginal progesterone administration, fetal genetic abnormalities, and recreational drug abuse. Participants answered a detailed questionnaire regarding their health, diet and lifestyle during this pregnancy. Clinical data was also collected, including gestational age, maternal age, parity, infant sex, indications for caesarean section (Table 8.1), and previous obstetric history. Maternal and fetal characteristics are summarised in Table 8.2.

Table 8.1: Indication for caesarean delivery in this cohort. In many cases there were multiple indications for caesarean delivery.

<table>
<thead>
<tr>
<th>Indication for caesarean section</th>
<th>Number of cases</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous caesarean section</td>
<td>35</td>
</tr>
<tr>
<td>Maternal diabetes</td>
<td>12</td>
</tr>
<tr>
<td>Elevated maternal BMI</td>
<td>5</td>
</tr>
<tr>
<td>Intra-uterine growth restriction</td>
<td>3</td>
</tr>
<tr>
<td>Tubal ligation</td>
<td>3</td>
</tr>
<tr>
<td>Tokophobia or previous traumatic vaginal delivery</td>
<td>3</td>
</tr>
<tr>
<td>Advanced maternal age</td>
<td>3</td>
</tr>
<tr>
<td>Macrosomic fetus</td>
<td>2</td>
</tr>
<tr>
<td>Breech position</td>
<td>2</td>
</tr>
<tr>
<td>Placenta previa</td>
<td>2</td>
</tr>
<tr>
<td>Maternal congenital heart defect</td>
<td>1</td>
</tr>
<tr>
<td>Placenta accreta</td>
<td>1</td>
</tr>
<tr>
<td>Previous myomectomy</td>
<td>1</td>
</tr>
<tr>
<td>Colitis</td>
<td>1</td>
</tr>
</tbody>
</table>
### TABLE 8.2: Maternal and fetal characteristics. Data are presented as median (IQR [Range]) or N (%). *P*-values represent analysis of inflammation present (+) vs inflammation absent (-). Abbreviations: N, number; n.s., not significant (*P* > 0.05); GA, gestational age; wk, weeks; IQR, interquartile range.

<table>
<thead>
<tr>
<th></th>
<th>All (n = 44)</th>
<th>Inflammation + (n = 21)</th>
<th>Inflammation - (n = 23)</th>
<th><em>P</em>-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>33 (29-35 [20-44])</td>
<td>33 (29-36 [20-42])</td>
<td>33 (29-34 [22-44])</td>
<td>n.s.</td>
</tr>
<tr>
<td>Gravidity</td>
<td>2 (2-3 [1-11])</td>
<td>2 (2-4 [1-6])</td>
<td>2 (2-3 [1-11])</td>
<td>n.s.</td>
</tr>
<tr>
<td>Parity</td>
<td>1 (1-2 [0-7])</td>
<td>1 (1-1 [0-5])</td>
<td>1 (1-2 [0-7])</td>
<td>n.s.</td>
</tr>
<tr>
<td>Maternal smoking</td>
<td>4 (9%)</td>
<td>1 (5%)</td>
<td>3 (13%)</td>
<td>n.s.</td>
</tr>
<tr>
<td>Pre-pregnancy BMI</td>
<td>28 (24-36 [18-52])</td>
<td>27 (23-33 [18-52])</td>
<td>30 (24-39 [18-49])</td>
<td>n.s.</td>
</tr>
<tr>
<td>Maternal diabetes</td>
<td>18 (41%)</td>
<td>7 (33%)</td>
<td>11 (48%)</td>
<td>n.s.</td>
</tr>
<tr>
<td>GA at birth (wk)</td>
<td>38 (38-39 [34-42])</td>
<td>38 (38-39 [36-41])</td>
<td>38 (37-39 [34-42])</td>
<td>n.s.</td>
</tr>
<tr>
<td>Male fetus</td>
<td>25 (57%)</td>
<td>11 (52%)</td>
<td>14 (61%)</td>
<td>n.s.</td>
</tr>
</tbody>
</table>

**Collection of cord blood**

Umbilical cord blood was collected by obstetric surgeons prior to delivery of the placenta. After birth, the cord was clamped and cut. The clamp was removed and cord blood was allowed to drip into a sterile collecting pot, then immediately handed off to the study investigator. Approximately 4 mL of cord blood was then poured into a BD Vacutainer® heparinised blood collection tube and gently mixed by inverting 8-10 times. All cord blood samples were processed within 20 minutes of collection.

**Immune stimulation**

Innate immune phentotyping was performed using a well-established, robust, validated, and quality-controlled protocol from Tobias Kollmann’s lab (University of British Columbia, Canada) [345-352]. Cord blood samples were plated out in duplicate in 200 µL aliquots onto custom made immune stimulation plates (plates manufactured by
Each well of each row was coated with one of eight nominated immune stimuli or controls. The stimuli used were selected to cover a broad range of immune pathways (Table 8.3). Resiquimod (R848), a powerful immune activator, was used as a toll-like receptor (TLR) 7/8 agonist. Lipopolysaccharide (LPS), a component of Gram negative bacterial cell walls, was used as a myeloid differentiation primary response 88 (Myd88) & TIR-domain-containing adapter-inducing interferon-β (TRIF) agonist. Peptidoglycan (PGN), the major outer membrane component of Gram positive bacterial cell walls, was used a TLR2 & nucleotide-binding oligomerisation domain (NOD) 1/2 agonist. Polyinosinic–polycytidylic acid (poly (I:C)), was used as a TLR3 agonist to simulate viral infection. Cyclic guanosine monophosphate–adenosine monophosphate (cGAMP) was used as a stimulator of interferon genes (STING) agonist to simulate intracellular pathogens. 5’ triphosphate double-stranded RNA (5’ppp-dsRNA) was used as a retinoic acid-inducible gene I (RIG-1) agonist to simulate viral infection. RPMI 1640 was used as a vehicle control for R848, LPS, PGN, and poly (I:C). LyoVec was used as a vehicle control for cGAMP and 5’ppp-dsRNA.

Samples were incubated in the sealed immune stimulation plates for 24 h at 37°C. After incubation, the plates were centrifuged at 4°C and the supernatant was transferred to a fresh microplate and frozen at -80°C until analysis.
TABLE 8.3: Details of immune stimulants used in this study.

<table>
<thead>
<tr>
<th>Stimulant</th>
<th>Abbreviation</th>
<th>Concentration</th>
<th>Brief description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unstimulated RPMI 1640 control</td>
<td>US-R</td>
<td>RPMI 1640 media only</td>
<td>Vehicle control used to compare to R848, LPS, PGN, and poly (I:C).</td>
</tr>
<tr>
<td>Resiquimod</td>
<td>R848</td>
<td>10 mM in RPMI 1640</td>
<td>TLR7/TLR8 agonist</td>
</tr>
<tr>
<td>Lipopolysaccharide</td>
<td>LPS (from <em>E. coli</em>)</td>
<td>100 ng/mL in RPMI 1640</td>
<td>Myd88 &amp; TRIF agonist</td>
</tr>
<tr>
<td>Peptidoglycan</td>
<td>PGN (from <em>S. aureus</em>)</td>
<td>10 μg/mL in RPMI 1640</td>
<td>TLR2 &amp; NOD1/2 agonist</td>
</tr>
<tr>
<td>Polyninosinic–polycytidylic acid</td>
<td>Poly (I:C)</td>
<td>100 μg/mL in RPMI 1640</td>
<td>TLR3 agonist</td>
</tr>
<tr>
<td>Unstimulated LyoVec control</td>
<td>US-L</td>
<td>LyoVec transfection reagent only</td>
<td>Vehicle control used to compare to cGAMP and 5’ppp-dsRNA</td>
</tr>
<tr>
<td>Cyclic guanosine monophosphate–adenosine monophosphate</td>
<td>cGAMP</td>
<td>20 μg/mL in LyoVec</td>
<td>STING agonist</td>
</tr>
<tr>
<td>5’ triphosphate double-stranded RNA</td>
<td>5’ppp-dsRNA</td>
<td>2 ng/mL in LyoVec</td>
<td>RIG-I agonist</td>
</tr>
</tbody>
</table>

**Quantitation of immune markers**

Human granulocyte-colony stimulating factor (G-CSF), interferon gamma (IFN-γ), interleukin 1 beta (IL-1β), interleukin 6 (IL-6), interleukin 8 (IL-8), interleukin 10 (IL-10) and tumour necrosis factor alpha (TNF-α) were measured by multiplex assay (R&D Systems Inc.) on a MAGPIX instrument (Luminex Corp) as per the manufacturer’s instructions. These factors were chosen for investigation as they are well characterised immune mediators that are relevant to the pathways stimulated in this study. All samples were diluted 1:2 prior to analysis.

**Placental histology**

Placentae and full thickness placental membranes were transported to the hospital’s histology department for routine examination by a senior histopathology scientist. The histology reports generated were used to classify inflammatory lesions of the placentae.
Standard culture-based screening for *Listeria* spp., *Ureaplasma* spp., and *Mycoplasma* spp. was also performed.

**Amniotic fluid cytokine analysis**

Approximately 10 mL of fluid was drawn into a syringe immediately following amniotomy, then transferred to sterile tubes and centrifuged at 40,000 X g at 4°C for 6 minutes to pellet. Levels of IL-6, IL-10, CXCL10, and G-CSF were measured in amniotic fluid supernatants by multiplex assay (R&D Systems Inc.) on a MAGPIX instrument (Luminex Corp) as per the manufacturer’s instructions. All samples were diluted 1:2 prior to analysis.

**Statistical analysis**

Comparisons were made between variables in placentae with VUE, CAM, FIR, and no inflammation (NI) using the Kruskal-Wallis test. Differences between groups were further analysed using the Mann-Whitney test. Presence of inflammatory reactions in the fetal blood vessels (either cord vasculitis or inflammation of the fetal vessels in the chorionic plate) was classified as FIR. Cytokine concentrations were corrected for dilution (x2) then expressed as a value relative to the appropriate control. SPSS (version 20.0, IBM SPSS) statistical software was used for data analysis. *P*-values < 0.05 were considered statistically significant.

**Results & discussion**

*Low grade placental inflammation is common in full-term pregnancies*

Of the 44 patients enrolled in this study 21 (47.7%) showed histological evidence of placental inflammation. This figure is similar to that reported by Romero *et al.* for women delivering by elective caesarean section [339]. Of those placentae positive for inflammation, 4 had DEC (of which we were only able to obtain cord blood from 2), 9 had VUE, 9 had CAM, and 12 had FIR (9 instances of umbilical cord vasculitis, 8 instances of inflammation in the fetal vessels of the chorionic plate). It is interesting to note that cases of VUE and CAM did not overlap in this study, suggesting separate
aetiologies. There was a high level of overlap between placentae with CAM and those with FIR. In most cases the inflammation was mild, with only four cases of high grade inflammation observed (all cases of VUE). Histology results are described in full in Table 8.4.

A number of maternal and fetal characteristics could contribute to placental inflammation. Obesity and cigarette smoking are known to cause systemic inflammation [353, 354]. In the present study the mean pre-pregnancy BMI of participants was elevated (30.1); however, there were no significant differences in pre-pregnancy BMI values between groups. The incidence of maternal smoking was actually higher in the inflammation negative group than the inflammation positive group (13% vs 5%). Diabetes could also contribute to placental inflammation; however, in the current study the incidence of maternal diabetes was higher in the inflammation negative group than in the inflammation positive group (48% vs 33%). Previous studies have demonstrated that placental immune function is partially sex-specific, and that the placenta responds to maternal inflammation in a sex-specific manner, with heightened responses in male placentae [355]. However, there was no significant difference in the rates of male births between groups in our study.

Routine hospital screening for *Ureaplasma* spp., *Mycoplasma* spp., and *Listeria* spp. was negative for all placentae; however, this is not surprising given that the presence of these bacteria in the amniotic cavity is normally associated with early preterm birth [356].

<table>
<thead>
<tr>
<th>Study ID</th>
<th>DEC</th>
<th>VUE</th>
<th>CAM</th>
<th>FIR</th>
<th>Other</th>
</tr>
</thead>
<tbody>
<tr>
<td>12</td>
<td>Focal, chronic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>48</td>
<td>Chronic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>40</td>
<td>Chronic</td>
<td>Chronic, high grade</td>
<td>Patchy, high grade, chronic</td>
<td>Early FIR of the fetal vessels in the chorionic plate</td>
<td>Meconium laden macrophages in placental membranes</td>
</tr>
<tr>
<td>45</td>
<td>Chronic</td>
<td>Focal, low grade</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>02</td>
<td>Focal, low grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>08</td>
<td>Low grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>32</td>
<td>Focal, low grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38</td>
<td>Focal, low grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39</td>
<td>High grade</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>47</td>
<td>Multifocal, low grade, chronic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>50</td>
<td>Patchy, high grade, chronic</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>13</td>
<td>Acute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23</td>
<td>Acute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26</td>
<td>Early/evolving, acute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>28</td>
<td>Acute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>44</td>
<td>Acute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>49</td>
<td>Early</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>Early, acute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>41</td>
<td>Early/evolving, acute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>43</td>
<td>Early/evolving, acute</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>31</td>
<td>Focal, early cord vasculitis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>21</td>
<td>Focal, acute cord vasculitis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>19</td>
<td>Focal, early cord vasculitis</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

142
CAM and FIR are associated with increased amniotic fluid IL-6 levels

Amniotic fluid cytokine levels were within previously-reported ranges in all cases (Figure 8.1). Amniotic fluid IL-6 was significantly elevated in pregnancies with CAM and FIR compared to those with no inflammation ($P = 0.0475$ and $P = 0.0322$, respectively). An amniotic fluid IL-6 level of >745 pg/mL is diagnostic of microbial invasion of the amniotic cavity and intra-amniotic inflammation [357]. Patients with CAM in the present study had a mean amniotic fluid IL-6 level of 765.7 pg/mL, while those with FIR had a mean amniotic fluid IL-6 level of 805.4 pg/mL. Similarly, an amniotic fluid G-CSF level of >2000 pg/mL is diagnostic of CAM [358]. Patients with CAM in the present study met this criterion, with a mean amniotic fluid G-CSF level of 2027.7 pg/mL. Those with FIR had a mean amniotic fluid G-CSF concentration of 2184.4 pg/mL. There were no significant differences between the groups with respect to concentrations of IL-10 (an anti-inflammatory cytokine) or CXCL10 (a putative marker of chronic chorioamnionitis arising from maternal fetal allograft rejection).

**FIGURE 8.1:** Amniotic fluid cytokine levels of patients with deciduitis (DEC, n = 3), vasculitis of unknown aetiology (VUE, n = 6), chorioamnionitis (CAM, n = 9), inflammatory reaction of the fetal vessels (FIR, n = 11), or no inflammation (NI, n = 20). Data are mean ± SEM. * $P < 0.05$ compared to the NI group.
**Cord blood immune response patterns differ with placental inflammation**

With the exception of IL-8, baseline expression of all cord blood cytokines was low (Table 8.5), as has previously been reported elsewhere [359, 360]. All forms of immune stimulation were associated with increased cord blood cytokine levels (Table 8.6). R848, LPS, and PGN elicited a strong inflammatory response across all cytokines, while poly (I:C), cGAMP, and 5’ppp-dsRNA elicited a weaker inflammatory response.

No significant differences were observed between groups after immune stimulation with R848, LPS, or PGN, possibly suggesting that fetal TLR7/8, NOD1/2, TLR2, Myd88, and TRIF pathways are not affected by placental inflammation.

Cord blood derived from fetuses exposed to CAM had significantly lower IL-8, IL-10, IL-1β, and G-CSF responses to poly (I:C) than those with no placental inflammation (IL-8 \( P = 0.0179 \), IL-10 \( P = 0.0262 \), IL-1β \( P = 0.0268 \), G-CSF \( P = 0.00135 \)) (Figures 8.2-8.5). Poly (I:C) is a ligand of TLR3 capable of activating downstream NF-κB and interferon regulatory factors 3 and 7 (IRF3/7), stimulating the production of pro-inflammatory cytokines, interferons and IL-10 [361]. Previous studies have revealed that cord blood natural killer (NK) cells have deficient TLR3 expression associated with an inability to respond to poly (I:C) [362]. Our data indicate that this dampened response to poly (I:C) is exaggerated in fetuses that are exposed to CAM during pregnancy. This is indicative of a greater impairment in NK TLR3 expression, which could result in heightened vulnerability to viral infections postnatally.

Immune responses in cord blood taken from fetuses with VUE did not differ from those of fetuses that were not exposed to inflammation. Fetuses whose placentae showed evidence of VUE had lower IL-10 responses to poly (I:C) than those not exposed to inflammation; however, this did not quite reach statistical significance \( (P = 0.0505) \) (Figure 8.3). Interestingly, fetuses with VUE had significantly higher IFN-γ responses to cGAMP than those that were exposed to CAM \( (P = 0.0485, \text{Figure 8.6}) \).
TABLE 8.5: Baseline levels of cytokines (pg/mL) in cord blood of patients with deciduitis (DEC, n = 2), vasculitis of unknown aetiology (VUE, n = 6), chorioamnionitis (CAM, n = 9), inflammatory reaction of the fetal vessels (FIR, n = 11), or no inflammation (NI, n = 20). Data are median ± IQR.

<table>
<thead>
<tr>
<th>Cytokine</th>
<th>Concentration in unstimulated RPMI 1640 control (pg/mL)</th>
<th>Concentration in unstimulated LyoVec control (pg/mL)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>DEC</td>
<td>VUE</td>
</tr>
<tr>
<td>G-CSF</td>
<td>32.0 ± 18.7</td>
<td>42.4 ± 49.5</td>
</tr>
<tr>
<td>IFN-γ</td>
<td>12.8 ± 12.8</td>
<td>8.8 ± 25.9</td>
</tr>
<tr>
<td>IL-1β</td>
<td>45.0 ± 40.0</td>
<td>22.9 ± 30.6</td>
</tr>
<tr>
<td>IL-6</td>
<td>54.4 ± 17.9</td>
<td>90.4 ± 378.9</td>
</tr>
<tr>
<td>IL-8</td>
<td>2797.6 ± 1872.8</td>
<td>1304.5 ± 3453.3</td>
</tr>
<tr>
<td>IL-10</td>
<td>4.6 ± 4.6</td>
<td>0.0 ± 2.8</td>
</tr>
<tr>
<td>TNF-α</td>
<td>16.5 ± 14.3</td>
<td>9.1 ± 7.3</td>
</tr>
</tbody>
</table>
TABLE 8.6: Mean fold change in cord blood cytokine levels following 24 hour incubation with various immune stimuli.

<table>
<thead>
<tr>
<th></th>
<th>R848</th>
<th>LPS</th>
<th>PGN</th>
<th>Poly (I:C)</th>
<th>CGAMP</th>
<th>5’ppp-dsRNA</th>
</tr>
</thead>
<tbody>
<tr>
<td>G-CSF</td>
<td>326.5</td>
<td>598.7</td>
<td>1126.1</td>
<td>17.7</td>
<td>13.8</td>
<td>11.6</td>
</tr>
<tr>
<td>IFN-γ</td>
<td>339.6</td>
<td>84.8</td>
<td>77.0</td>
<td>6.5</td>
<td>60.8</td>
<td>18.0</td>
</tr>
<tr>
<td>IL-1β</td>
<td>3948.7</td>
<td>2062.7</td>
<td>1157.0</td>
<td>25.1</td>
<td>225.7</td>
<td>163.6</td>
</tr>
<tr>
<td>IL-6</td>
<td>847.5</td>
<td>855.1</td>
<td>870.1</td>
<td>172.8</td>
<td>232.2</td>
<td>331.3</td>
</tr>
<tr>
<td>IL-8</td>
<td>96.9</td>
<td>92.4</td>
<td>105.6</td>
<td>91.0</td>
<td>5.5</td>
<td>9.7</td>
</tr>
<tr>
<td>IL-10</td>
<td>523.4</td>
<td>455.3</td>
<td>139.7</td>
<td>17.2</td>
<td>3.1</td>
<td>12.2</td>
</tr>
<tr>
<td>TNF-α</td>
<td>3507.0</td>
<td>610.5</td>
<td>748.0</td>
<td>10.4</td>
<td>178.6</td>
<td>103.8</td>
</tr>
</tbody>
</table>

Cord blood taken from fetuses with FIR had significantly lower IL-8, IL-10, and G-CSF responses to poly (I:C) than those with no placental inflammation (IL-8 \( P = 0.0427 \), IL-10 \( P = 0.0262 \), G-CSF \( P = 0.0132 \)) (Figures 8.2, 8.3, 8.5). Unsurprisingly, given the overlap in cases of CAM and FIR, these responses were very similar to those seen in the CAM group. Interestingly, the FIR group differed from the CAM group in its response to cGAMP and 5’ppp-dsRNA. Stimulation with cGAMP elicited a significantly higher IL-6, IL-8, and IL-1β response in fetuses with FIR than those not exposed to inflammation (IL-6 \( P = 0.00695 \), IL-8 \( P = 0.0336 \), IL-1β \( P = 0.0475 \)) (Figures 8.2, 8.4, 8.7). Similarly, stimulation with 5’ppp-dsRNA elicited a significantly higher IL-6 and IL-8 response in fetuses with FIR than those not exposed to inflammation (IL-6 \( P = 0.0228 \), IL-8 \( P = 0.0294 \)) (Figures 8.2 & 8.7). cGAMP and 5’ppp-dsRNA are ligands of STING and RIG-I, respectively. Both STING and RIG-I drive downstream type 1 interferon production, NF-κB activation, and antiviral gene transcription to control viral infections [363, 364]. There exists much cross-talk and interplay between the STING and RIG-I pathways, so it is not surprising that their stimulation resulted in similar reactions between groups in this study.
FIGURE 8.2: IL-8 levels in fetal cord blood plasma relative to unstimulated controls. Data is stratified according to type of placental inflammation (DEC n = 2, VUE n = 7, CAM n = 9, FIR n = 12, NI n = 23). * P < 0.05.
FIGURE 8.3: IL-10 levels in fetal cord blood plasma relative to unstimulated controls. Data is stratified according to type of placental inflammation (DEC n = 2, VUE n = 7, CAM n = 7, FIR n = 7, NI n = 19). * P < 0.05.
FIGURE 8.4: IL-1β levels in fetal cord blood plasma relative to unstimulated controls. Data is stratified according to type of placental inflammation (DEC n = 2, VUE n = 7, CAM n = 9, FIR n = 12, NI n = 23). * P < 0.05.
FIGURE 8.5: G-CSF levels in fetal cord blood plasma relative to unstimulated controls. Data is stratified according to type of placental inflammation (DEC n = 2, VUE n = 7, CAM n = 9, FIR n = 12, NI n = 23). * $P < 0.05$, ** $P < 0.01$. 
FIGURE 8.6: IFN-γ levels in fetal cord blood plasma relative to unstimulated controls. Data is stratified according to type of placental inflammation (DEC n = 1, VUE n = 7, CAM n = 9, FIR n = 11, NI n = 20). * P < 0.05.
FIGURE 8.7: IL-6 levels in fetal cord blood plasma relative to unstimulated controls. Data is stratified according to type of placental inflammation (DEC n = 2, VUE n = 7, CAM n = 9, FIR n = 12, NI n = 23). * P < 0.05, ** P < 0.01.
Fetuses exposed to CAM and FIR showed significantly reduced immune responses to stimulation with poly (I:C) and significantly increased responses to stimulation with cGAMP and 5’ppp-dsRNA. Given that TLR3 activates transcription of IRF3, and that STING and RIG-I activate transcription of IRF3 independently of the action of TLR3, this opposing reaction may be a compensatory mechanism. Interestingly, these aberrant reactions were almost exclusive to fetuses exposed to CAM and FIR and not observed in fetuses exposed to VUE. The significance of such a finding is unclear, but could be related to the differences in aetiology of the placental lesions.

A potential link with the intra-amniotic microbiome

It is unclear whether the inflammation seen in this cohort has a microbiological aetiology or a sterile aetiology. Recent research has suggested that the intra-amniotic space is, in some pregnancies, colonised by bacteria prior to birth [210], and that this colonisation may alter fetal immune development [18, 213]. Differences in placental inflammation and neonatal immune responses may, therefore, be reflective of differences in prenatal microbial colonisation.

The bacterial profiles of the amniotic fluid and meconium of the patients enrolled in this study were investigated and discussed in Chapter 5 of this thesis. Here we have sought to identify whether the bacterial 16S rRNA gene profiles of these samples might explain the inflammatory and immune response results reported here. Intriguingly, there was no relationship between the amniotic fluid microbiota (in terms of alpha or beta diversity, or individual species abundance) and fetal immune responses or placental inflammation. Nor could we identify any relationship between the meconium microbiota and placental inflammation. However, we did find that neonates whose meconium was dominated by reads that mapped to Pelomonas puraquae (i.e. the number of P. puraquae reads was greater than the sum of all other reads) had increased cord blood TNF-α, IL-6 and IL-1β responses to stimulation with R848, and increased IL-6 and IL-1β responses to stimulation with both LPS and PGN (all $P < 0.02$). Whether these differences demonstrate immune sensitisation or immune immaturity in this context is unclear.
Given that LPS is the major outer membrane component of Gram negative bacterial cell walls and that PGN is the major outer membrane component of Gram positive bacterial cell walls, responses to these stimulants might differ depending on whether the fetus has been exposed to Gram positive or negative bacteria. Therefore, we divided the participants in this study into groups depending on whether their meconium or amniotic fluid samples contained Gram negatives only, Gram positives only, or a combination of both. There were no meconium samples which contained Gram positives only, so for meconium samples comparisons were only made between those containing Gram negatives only and those containing a combination of Gram positive and negative bacteria. In all cases stimulation with LPS and PGN elicited a strong immune response. Surprisingly, cord blood immune responses to LPS and PGN did not vary based on the presence of Gram positive or Gram negative bacterial DNA in the amniotic fluid or meconium (Table 8.7). This result may indicate that the presence of these bacteria in the intra-amniotic space do not influence the prenatal immune system. However, the number of subjects for whom both microbiome and immune response data was available was quite low (n = 36 for meconium samples, n = 29 for amniotic fluid samples), and there was a high level of inter-subject variation in immune responses to LPS and PGN. Therefore, we may not have had sufficient power to detect prenatal immune conditioning to Gram positive or negative bacteria in this study. It is also possible that we have detected cell-free bacterial DNA, meaning that these infants were not exposed to LPS or PGN from bacterial cell membranes prior to birth. However, we have previously demonstrated that first pass meconium contains bacteria with intact cell walls (Chapter 7). Additionally, DNA was extracted from pelleted cells following centrifugation, which is unlikely to have pelleted cell-free DNA.

It is possible that at least some of the placental inflammation found in this cohort is non-bacterial in origin. Previous studies have found sterile placental inflammation to be more common than microbial-associated placental inflammation in women giving birth preterm with intact membranes [365, 366]. However, inflammation that has been described as ostensibly “sterile” may not have a sterile aetiology at all. It is possible that sub-clinical viral infections are responsible for some forms of placental inflammation seen in the present study and in previous studies of “sterile” inflammation. Viruses are commonly overlooked in microbiome studies as they are more difficult to profile than bacteria or fungi; largely due to the absence of a universal viral gene that can be used
for easy taxonomic profiling of the majority of viral taxa. To date, the presence and effect of the intra-amniotic virome has not been thoroughly explored. Lim et al. reported the presence of a diverse range of bacteriophages and a small number of eukaryotic viruses in amniotic fluid samples from healthy, full-term pregnancies [62]. Our data indicate that the presence of CAM and FIR is associated with decreased immune responses to a TLR3 agonist, which is indicative of a decreased viral immune response. Therefore, viral colonisation of the fetal compartment during pregnancy may be the underlying cause of these inflammatory lesions and the altered immune responses observed in the present study.
TABLE 8.7: Mean cord blood cytokine levels after stimulation with LPS or PGN in neonates whose meconium contained Gram negative bacteria only or a combination of Gram negative and positive bacteria, and in neonates whose amniotic fluid contained Gram negative bacteria only, Gram positive bacteria only, or a combination of Gram negative and positive bacteria.

<table>
<thead>
<tr>
<th>Meconium</th>
<th>Cytokine</th>
<th>Gram – only</th>
<th>Gram + only</th>
<th>Combination</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulant</td>
<td></td>
<td>n = 29</td>
<td>n = 0</td>
<td>n = 7</td>
<td></td>
</tr>
<tr>
<td>LPS</td>
<td>TNF-α</td>
<td>630.28</td>
<td>226.29</td>
<td>0.18</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-6</td>
<td>1199.22</td>
<td>297.81</td>
<td>0.25</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-8</td>
<td>76.97</td>
<td>343.34</td>
<td>0.58</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-10</td>
<td>695.57</td>
<td>127.26</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-1β</td>
<td>2528.96</td>
<td>421.60</td>
<td>0.07</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IFN-γ</td>
<td>19.50</td>
<td>23.46</td>
<td>0.23</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G-CSF</td>
<td>804.07</td>
<td>342.10</td>
<td>0.32</td>
<td></td>
</tr>
<tr>
<td>PGN</td>
<td>TNF-α</td>
<td>818.06</td>
<td>189.11</td>
<td>0.10</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-6</td>
<td>1214.78</td>
<td>328.52</td>
<td>0.26</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-8</td>
<td>100.42</td>
<td>345.08</td>
<td>0.66</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-10</td>
<td>190.94</td>
<td>99.56</td>
<td>0.38</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IL-1β</td>
<td>1393.81</td>
<td>413.42</td>
<td>0.17</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IFN-γ</td>
<td>9.98</td>
<td>16.33</td>
<td>0.20</td>
<td></td>
</tr>
<tr>
<td></td>
<td>G-CSF</td>
<td>1431.61</td>
<td>797.11</td>
<td>0.28</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Amniotic fluid</th>
<th>Cytokine</th>
<th>Gram – only</th>
<th>Gram + only</th>
<th>Combination</th>
<th>P</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stimulant</td>
<td></td>
<td>n = 9</td>
<td>n = 9</td>
<td>n = 11</td>
<td></td>
</tr>
<tr>
<td>LPS</td>
<td>TNF-α</td>
<td>862.71</td>
<td>609.93</td>
<td>436.87</td>
<td>0.68</td>
</tr>
<tr>
<td></td>
<td>IL-6</td>
<td>791.65</td>
<td>1443.16</td>
<td>550.11</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>IL-8</td>
<td>34.29</td>
<td>37.72</td>
<td>162.05</td>
<td>0.45</td>
</tr>
<tr>
<td></td>
<td>IL-10</td>
<td>1036.38</td>
<td>275.19</td>
<td>139.59</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>IL-1β</td>
<td>1582.00</td>
<td>2736.66</td>
<td>1781.54</td>
<td>0.77</td>
</tr>
<tr>
<td></td>
<td>IFN-γ</td>
<td>22.03</td>
<td>21.22</td>
<td>16.99</td>
<td>0.48</td>
</tr>
<tr>
<td></td>
<td>G-CSF</td>
<td>1220.42</td>
<td>378.60</td>
<td>592.87</td>
<td>0.87</td>
</tr>
<tr>
<td>PGN</td>
<td>TNF-α</td>
<td>1028.72</td>
<td>640.72</td>
<td>303.05</td>
<td>0.53</td>
</tr>
<tr>
<td></td>
<td>IL-6</td>
<td>794.08</td>
<td>1481.30</td>
<td>564.34</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>IL-8</td>
<td>35.43</td>
<td>52.12</td>
<td>210.63</td>
<td>0.65</td>
</tr>
<tr>
<td></td>
<td>IL-10</td>
<td>289.17</td>
<td>101.80</td>
<td>69.31</td>
<td>0.23</td>
</tr>
<tr>
<td></td>
<td>IL-1β</td>
<td>962.04</td>
<td>1635.26</td>
<td>734.72</td>
<td>0.57</td>
</tr>
<tr>
<td></td>
<td>IFN-γ</td>
<td>9.84</td>
<td>13.43</td>
<td>7.96</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>G-CSF</td>
<td>2211.91</td>
<td>767.33</td>
<td>955.48</td>
<td>0.98</td>
</tr>
</tbody>
</table>

Summary

Exposure to inflammation in-utero is associated with altered immune reactions at birth. Different types of inflammatory lesions were associated with distinct fetal immune response profiles. Exposure to bacteria/bacterial DNA prior to birth does not appear to underpin this effect. The immunological significance of these observations remains to be determined.
“So we have re-told the birth narrative from a holobiont perspective... In this new narrative, we see ‘birth’ as involving the reproduction of the holobiont. In other words, both the mammal and her persistent microbial populations have to be reproduced.”

Overview

This thesis has explored key questions relating to the establishment of the human microbiome. In particular, I have sought to use stringent, high quality methods to determine the extent to which the human fetus is exposed to bacteria and the metabolomic and immunological correlates of this exposure.

This thesis was structured around five overarching research aims:

**Aim 1:** To develop and validate improved methodologies for studying the fetal microbiome to create more robust and reliable data in this field.

**Aim 2:** To characterise the bacterial composition of meconium and amniotic fluid at term in normal pregnancies.

**Aim 3:** To examine the relationship between maternal health, fetal bacterial colonisation, placental inflammation and fetal immune development.

**Aim 4:** To ascertain whether the apparent fetal gut microbiome represents living, metabolically active cells, or dead cells and bacterial products transferred from the mother.

**Aim 5:** To determine whether amniotic fluid contains bacterial DNA in mid-pregnancy and, if so, to characterise the taxonomic profile of these samples.

In this chapter I will draw together key findings presented in each of the preceding chapters to address these aims. I will also highlight the limitations of these studies and avenues for future research, and discuss the broader implications of these findings.

**Aim 1: To develop and validate improved methodologies for studying the fetal microbiome to create more robust and reliable data in this field.**

Over the past decade increasing interest has been paid to the purported fetal microbiome. NGS studies of the amniotic fluid and fetal gut have led to the hypothesis that the fetus is colonised with bacteria prior to birth. However, this is a field with
numerous technical challenges. In particular, there is a very low abundance of bacteria in the intra-amniotic space, allowing external contamination to heavily confound the resulting data. Further, studies in this field differ widely in their use of extraction kits, PCR primers, sequencing platforms, and bioinformatics tools. This makes a direct comparison of results between such studies very difficult. The first aim of this thesis was, therefore, to develop and validate improved methodologies for studying the fetal microbiome.

This section of the thesis has implications beyond the study of the perinatal microbiome. In Chapter 3 I have demonstrated that choice of extraction kit has a marked influence on both the ability to extract DNA from meconium and on the resulting microbial profile. Others have demonstrated this phenomenon in other samples types [367-371]. Researchers should refer to such studies when selecting an extraction method for their own sample type, or, if no such studies exist, perform their own comparisons. Importantly, given the results presented in Chapter 3, studies of the meconium microbiome that use different extraction kits are not comparable.

In Chapter 4 I have shown that the primary source of laboratory contamination in my microbiome workflow was not the “kitome”, as previously thought, but rather the “mixome”. I was able to show that dsDNase treatment can eliminate this contamination, allowing low titre microbiomes to shine through beyond the level of background contamination. This kind of treatment is important for any study of a low biomass environment (such as human lungs, blood, and bladder), or any study where a high degree of sensitivity is necessary.

If the bacterial DNA sequencing performed in Chapter 5 were to have been done without the dsDNase treatment described in Chapter 4, the results would be difficult to interpret and might have led to false conclusions. Pelomonas puraquae was a major contaminant in untreated extraction controls and PCR controls tested in Chapter 4. Treatment of the PCR reagents with dsDNase reduced the levels of P. puraquae DNA in all controls drastically, completely eradicating it in some instances. P. puraquae DNA was consistently present in relatively high levels in our meconium samples (mean
reads: 3,437; interquartile range: 1,415-4,441), and absent from both PCR controls and 4/5 of the relevant extraction controls. These data suggest that the *P. puraquaee* DNA detected in our meconium samples was truly present, and not an artefact of contamination. Untreated extraction controls described in Chapter 4 contained an average of 25,913 reads that mapped to the *Pelomonas* genera, while untreated PCR controls contained an average of 20,891 reads that mapped to the *Pelomonas* genus. This background contamination is likely to have outweighed what was found in our samples, which might have led us to conclude that the observed *P. puraquaee* DNA in our samples originated from reagent-based contaminant. Conversely, our untreated extraction and PCR controls contained numerous reads belonging to *Escherichia/Shigella* and *Enterobacter*. Species of these genera, such as *Escherichia coli* and *Enterobacter cloacae*, are typical members of the adult gut microbiota. If we did not treat our PCR reagents prior to the work performed in Chapter 5, the resulting data might have indicated the presence of these bacteria in our samples, leading to inaccurate conclusions. However, when our meconium samples were processed using dsDNase-treated PCR reagents, we did not recover any reads belonging to these genera. As noted in Chapter 4, remnant DNA from the *E. coli* used to produce Taq DNA polymerase might be amplified by PCR and result in the false assumption that *Escherichia/Shigella* are present in a given sample. It is interesting to note that *Escherichia sp.* and *Enterobacter sp.* have been frequently reported as major parts of the meconium microbiome [11, 13, 24, 26, 55, 56, 58, 59, 63, 297]. Caution should be taken when interpreting these results, as our findings suggest that these bacteria might be reagent-based contaminants.

**Aim 2: To characterise the bacterial composition of meconium and amniotic fluid at term in normal pregnancies.**

This thesis presents the first study of the meconium and amniotic fluid microbiomes to use full-length 16S rRNA gene sequencing. This technology has allowed us to characterise these environments with a high level of resolution, in most instances to the species level. Meconium samples were strongly dominated by reads that mapped to *P. puraquaee* with a high level of sequence homology (99%). By carefully controlling for and actively removing reagent-based contamination, we are able to state that *P. puraquaee* DNA is likely to be truly present in these meconium samples. However, given
that it did appear in one meconium extraction control, we cannot completely rule out the possibility that it is an external contaminant. We are the first to report this species as a major pioneering species of the human fetal gastrointestinal tract, although previous studies have reported the detection of *Pelomonas* spp. DNA in amniotic fluid and endometrial swabs [14, 291, 292]. Other bacterial genera detected in the meconium samples studied here include *Staphylococcus, Propionibacterium, Corynebacterium, Rothia, Klebsiella, and Streptococcus*. Amniotic fluid samples were more diverse than meconium samples, and were largely dominated by reads that mapped to typical human skin commensals, including *Propionibacterium acnes, Staphylococcus haemolyticus,* and *Staphylococcus lugdunensis*. In support of our findings, *Staphylococcus* spp., *Propionibacterium acnes, Rothia* spp., *Corynebacterium* spp., *Klebsiella* spp., and *Streptococcus* spp. have previously been cultured from first-pass meconium and amniotic fluid samples [11, 58, 78]. A major point on which our results differ from those of culture studies of the first-pass meconium is the presence of *Enterococcus*. Both Moles *et al.* and Jimenez *et al.* found *Enterococcus* to be a major cultivable genus in meconium [58, 78]. We did not detect any sequences that mapped to *Enterococcus* in meconium nor amniotic fluid. The primers used for our PacBio sequencing covered 89.1% of *Enterococcus* spp.; however, there is still the possibility that key species were not covered with this primer set.

Meconium and amniotic fluid samples displayed a high level of inter-relatedness, with 32.7% of detected species being shared between the two sites, and 93.6% of total reads belonging to shared species. These data might, therefore, support a model of initial bacterial seeding into the amniotic fluid followed by dispersion into the fetal gut via swallowing. Interestingly, *P. puraquae* was detected in amniotic fluid samples, and appeared to expand in dominance in meconium samples. Environmental selection pressures, such as salt levels, pH, and pO$_2$, may therefore shape the meconium microbiome, allowing a subset of swallowed bacteria to survive and thrive. Dispersal from the amniotic fluid to the gut might be a selective process that determines the assembly of the fetal gut microbiome. This raises the question of whether microbial community assembly in the fetus is dominated by stochastic or deterministic processes.
It is important to interpret bacterial DNA sequencing results in the light of microbial ecology. As De Goffau et al. write: “many papers in the literature report microbial associations that go against basic understanding of microbial ecology, some of which can be likened to reporting blue whales in the Himalayas or African elephants in Antarctica” [240]. To avoid such a faux pas, we have attempted to assess our microbiological findings in the context of the environmental parameters of the fetal gut and amniotic fluid. Unfortunately, no data exists on the basic environmental parameters of meconium (such as salt levels, pH, and pO$_2$), so we are not able to assess our findings in this way. While experimental data on the oxygenation of the fetal gut does not exist, others have suggested that the fetal gut is an aerobic (albeit low-oxygen) environment which becomes anaerobic following birth [100, 372-374]. This assumption is based on the observation that the initial colonisers of the infant gut in the first days postpartum are facultative anaerobes, which are succeeded by strict anaerobes over the first weeks of life [13, 374]. Of the major OTUs identified in our survey of the fetal gut (those with >500 reads of >95% sequence homology across all meconium samples), 9/16 were aerobes, 6/16 were facultative anaerobes, and 1/16 was an aerotolerant anaerobe. These taxa are what we might expect to see as pioneering species of the fetal gut, based on the assumption that the above theory is correct. Interestingly, 6/16 of these species were motile, Gram negative, aerobes. Motility can confer a competitive advantage to a bacterial cell depending on its environment. However, it is unclear whether this is the case for bacterial pioneers of the fetal gut. The amniotic fluid is a low-oxygen environment (with pO$_2$ values previously reported as 2.8 mm Hg [375]), although both aerobes and anaerobes have previously been cultured from amniotic fluid taken from women with intact membranes [376]. The amniotic fluid samples investigated here were dominated by *P. puraquae*, *P. acnes*, various *Staphylococcus* spp., and various *Streptococcus* spp. Apart from *P. puraquae*, which is an aerobe, these other bacteria are facultative anaerobes. They are therefore suited to a low oxygen environment such as amniotic fluid.

The observed bacterial signatures in these samples do not necessarily imply the presence of a true “microbiome”. A microbiome is not simply a group of living microbes, but rather an ecological system of interacting organisms and their environment. Our results might reflect the presence of bacteria that are transferred from mother to fetus, but are not suited to survive within the womb. These bacteria may
therefore be transiently present, but not true colonisers. Indeed, this may explain the low abundance of bacteria in the womb. In Chapter 6 of this thesis we reported that bacteria may be present in the amniotic fluid of women with uncomplicated pregnancies as early as mid-gestation. If a true microbiome was established at this stage, we might expect it to become increasingly abundant after months of incubation in-utero. However, it remains as a low abundance microbiome that varies greatly between individuals at term. Another explanation for the low titre of bacteria found within the intra-amniotic space is the presence of both fetal and maternal immunological components in amniotic fluid which may suppress bacterial growth (for instance maternal and fetal neutrophils [377]). Alternatively, non-bacterial microbial elements such as bacteriophages may curate the bacterial microbiome in-utero and prevent it from expanding in biomass.

Together, the data presented in this thesis provide new evidence that bacteria and/or bacterial DNA can access the fetus prior to birth. This finding has potentially major implications for our understanding of fetal microbial and immune development.

**Aim 3: To examine the relationship between maternal health, fetal bacterial colonisation, placental inflammation and fetal immune development.**

We did not observe any major differences in meconium or amniotic fluid microbiomes based on maternal pre-pregnancy BMI, parity, diabetes, asthma, atopy, smoking, ethnicity, or gestational age at delivery. We may have failed to detect changes in the fetal microbiome based on maternal health parameters due to a lack of power. As Boers et al. pointed out, small cohorts are insufficient to detect changes in “normal” microbiota [298]. It would be expected that the maternal microbiome would vary based on factors such as diabetes status, BMI, and diet. However, we were not able to detect the effects of such variables within the amniotic space.

Fetal immune responses were altered in pregnancies with inflammation in the membranes and fetal tissue of the placenta. In these cases of feto-placental inflammation, amniotic fluid IL-6 levels were also raised. Placental inflammation was
associated with significantly lower levels of the microbial metabolite acetate in meconium, potentially suggesting a link between the fetal gut microbiota, placental inflammation, and immune responses. Surprisingly, there was no relationship between the amniotic fluid microbiota and fetal immune responses. However, we did find that neonates whose meconium was dominated by *P. puraquae* had increased TNF-α, IL-6 and IL-1β responses to stimulation with R848 (a TLR7/8 agonist), and increased IL-6 and IL-1β responses to stimulation with both LPS (a Myd88 and TRIF agonist) and PGN (a TLR2 and NOD1/2 agonist). Whether these differences demonstrate immune sensitisation or immune immaturity in this context is unclear. However, given that the meconium microbiota represent the first microbes that the intestinal immune system will encounter, this early exposure may play a role in preparing the immune system for life outside the womb.

**Aim 4: To ascertain whether the apparent fetal gut microbiome represents living, metabolically active cells, or dead cells and bacterial products transferred from the mother.**

Microbiome studies based on DNA sequencing frequently interpret the data as a reflection of living bacterial microbiota; however, this may not necessarily be the case. In Chapter 7 I have used molecular methods to demonstrate that there are viable bacterial cells in meconium. PMA allows selective amplification of DNA that originates from intact cells only. However, PMA is not a perfect tool, and is very difficult to optimise for use in complex samples [326]. Thus, the evidence provided in Chapter 7 is suggestive of the presence of viable bacterial cells in meconium, but is not conclusive. Further, viability does not necessarily translate to activity. Bacterial cells may be present within the amniotic cavity, but not in an active state. To conclusively determine if there are live, metabolically active bacteria in the intra-amniotic cavity, we would need to perform metatranscriptomics or culture-based experiments. However, organisms that are static *in-utero* may return to an exponential growth phase when given the required media. Therefore, even bacterial culture cannot accurately ascertain the true metabolic activity of fetal bacteria; it can, however, provide data on cell viability and pave the way for strain-level genomic analyses.
There is already a wealth of culture-based data on the meconium microbiome, dating back almost a century. In 1927 Burrage collected first-pass meconium from 100 healthy full-term infants, and was able to culture bacteria from 38% of these [43]. Hall and O’Toole performed similar experiments in 1934, on a cohort of 50 healthy full-term newborns, also finding that 38% of samples were colonised [44]. These authors reported that the majority of isolates belonged to the Micrococcus genera – a human skin commensal that is sometimes associated with pathogenic infections. In 1936 Snyder tested the sterility of the fetal gastrointestinal tract using rectal cannulation and intrarectal swabbing of 64 healthy newborns [45]. All samples were collected within 4 hours of birth, and 29 were collected within the first 30 minutes of life. Snyder reported that 6.3% of samples yielded positive results by direct culture, while 36% were positive by enrichment culture. This particular study was able to culture mainly skin commensals, which led the author to believe that the samples were contaminated by the perianal skin. However, the data presented in Chapter 5 of this thesis suggests that fetuses may be exposed to skin commensals in-utero. In 2008 Jiminez et al. were able to culture bacteria from 100% of the 21 samples they collected from healthy newborns within the first two hours of life [78]. These authors reported that the majority of isolates belonged to the genera Enterococcus and Staphylococcus. More recently, Moles et al. combined culture-based and molecular techniques to characterise the meconium microbiome of preterm neonates [58]. These authors were able to culture bacteria from 78.6% of these samples, also finding that Staphylococcus spp. were the predominant isolate. These studies support our finding that Staphylococcus spp. are common members of both the amniotic fluid and meconium microbiomes. While none of these culture-based studies identified P. puraquae, this is unsurprising as this bacterium was not described until 2007. However, Moles et al. were able to culture Pseudomonas in their samples [58, 78]. 16S rRNA gene sequences belonging to P. puraquae may have mapped to this taxon prior to the discovery of P. puraquae. In fact, Pelomonas saccharophila was originally classified as Pseudomonas saccharophila, but was reclassified in 2005 upon the creation of the Pelomonas genus. Overall, the weight of evidence from culture-based studies suggests that there are live bacteria present in the human gut at birth which can be subsequently cultured from meconium. The findings presented in this thesis support this hypothesis, while additionally suggesting that there are other, fastidious live bacteria present in meconium.
Our finding that short chain fatty acids (SCFAs) are present in meconium (Chapter 5) provides secondary evidence for the theory that there are living, metabolically active bacteria in the fetal gut, as SCFAs may be produced by bacteria within the intra-amniotic space. However, evidence from animal models suggests that SCFAs might be transferred from the maternal circulation to the fetus via the placenta [4], so the detection of SCFA in meconium does not conclusively prove bacterial metabolic activity. Interestingly, SCFA receptors (G protein-coupled receptors) have previously been identified in utero-placental tissue [378]. Expression of these receptors in fetal membranes increases in response to LPS, and is increased in fetal membranes from women with infection-mediated preterm birth [378]. Therefore, maternal- or fetal-derived SCFAs may play an important immunomodulatory role in-utero.

**Aim 5: To determine whether amniotic fluid contains bacterial DNA in mid-pregnancy and, if so, to characterise the taxonomic profile of these samples.**

In Chapter 6 I screened 1198 mid-gestation amniotic fluid samples (14-20 weeks) and found that 19.9% of them contained detectable levels of bacterial DNA. This finding provides evidence that the intra-amniotic microbiome may begin far earlier than previously expected. Only three other studies to date have used next generation sequencing to examine the mid-gestation amniotic fluid microbiome [14, 300, 302]. These studies have reported rates of mid-gestation amniotic fluid colonisation that range from 3-100%. However, true level of colonisation and the true composition of these samples are difficult to ascertain due to background contamination. Pathogenic infections of the amniotic fluid have also been detected early in pregnancy, from 20-22 weeks [155, 161]. However, overall, the timing of colonisation of the amniotic cavity in normal, healthy pregnancies remains inconclusive. Thus, the evidence provided in this thesis contributes to our understanding of the time-course of fetal microbiome seeding by providing evidence that bacterial DNA is present in amniotic fluid in mid-gestation in some pregnancies.

Transcriptome analysis of healthy human blood has revealed that it contains low levels of bacteria, with a microbial signature dominated by Proteobacteria [379]. Perez et al.
demonstrated that the peripheral blood mononuclear cells (PBMC) of lactating women contained a higher complexity of bacteria and higher level of inter-individual variation than those of non-pregnant, non-lactating controls [125]. Further, they showed a marked increase in live bacteria that could be cultured from the mesenteric lymph nodes of mice 5-6 days before birth that decreased drastically within 24 h of birth. Together this data may suggest that there is an upregulation of bacterial trafficking in maternal blood and lymph systems during pregnancy that increases towards the end of pregnancy. This may act to support breastmilk bacterial colonisation. However, it may also have the effect of introducing bacteria to the intra-amniotic space towards the end of gestation. Indeed, the data presented in this thesis indicate that the rate of bacterial colonisation of the amniotic fluid increases across gestation (from 19.9% at mid-gestation to 83.7% at full-gestation); a finding that might support the hypothesis that bacterial translocation increases over the course of pregnancy. However, it should be noted that these samples were not collected longitudinally from the same women, so our data is unable to directly address this hypothesis.

The mid-gestation and full-gestation amniotic fluid samples described in this thesis (Chapters 5 and 6) were processed using almost identical methods, providing a unique opportunity to compare such samples without the complications of inter-study variation. This might allow us to compare the two to ascertain whether the amniotic fluid microbiome is stable or dynamic across gestation. The maternal gut microbiome is known to change dramatically across gestation [12]. As it has been hypothesised that the fetus is seeded with bacteria from the maternal gut [210], it may also be hypothesised that the intra-amniotic microbiome changes throughout gestation. Indeed, the mid- and full-gestation amniotic fluid samples described in this thesis did differ in terms of species present; however, these differences might not be attributable to gestational age alone. The groups compared herein also differed considerably in their demographic characteristics and geographical location. The mid-gestation samples were taken from a Swedish cohort that consisted of almost exclusively Caucasian women with low BMIs who did not smoke during pregnancy. By comparison, the full-gestation samples were taken in Australia from a more ethnically heterogeneous group with higher BMIs and a higher rate of smoking during pregnancy. Additionally, all women in the full-gestation group were multiparous, while only 83% of women in the mid-gestation group were. These factors are important considerations, as each might
influence the maternal microbiome, and thereby shape the amniotic fluid microbiome. Cigarette smoking has a dysbiotic effect on the gut microbiome, potentially mediated through increased oxidative stress, alterations of intestinal tight junctions and mucin, and changes in acid-base balance [380]. It also has a dysbiotic effect on the vaginal microbiome via induction of vaginal _Lactobacillus_ phages [381]. Obesity has repeatedly been shown to influence and be influenced by the gut microbiome [382-384]. Additionally, both vaginal [385, 386] and gut microbiomes [282, 387-389] differ between ethnic groups. Although it has never been tested, the intra-uterine microbiome may change with pregnancy and birth, and thus, parity and gravida might be important considerations in this type of work.

Our results may also reflect geographic variation in amniotic fluid bacterial composition. The internal human microbiome is influenced by the external environmental microbiome in which the host lives [282, 284]. Local variations in food, water, plant, animal, soil, and built environment microbiomes shape variations in human microorganisms at the population level [283, 285]. _A. temperans_ was one of the most abundant reads in our Swedish mid-gestation samples. It was present in seven samples, and was the dominant read in two samples. All recorded cases of clinical isolation of _A. temperans_ have been in Europe (Sweden, France, and the UK) [313, 314] and most of these have been in Gotenberg, Sweden [313] – the very city that our mid-gestation samples were taken in. Similarly, _Pelomonas_ sp., which was one of the most abundant reads in our Australian full-gestation samples, has previously been isolated in Australian drinking water [286].

The data produced here indicates that there are key differences in the bacterial community composition of amniotic fluid across gestation. However, the cohorts from which these samples were derived differed significantly in both demographic and geographic factors. Therefore, the differences observed in these two cohorts might reflect differences in the maternal microbiome more generally between the two populations. Furthermore, 10 mL of amniotic fluid was used for DNA extraction in the full-gestation cohort, while only 1 mL was used for the mid-gestation cohort. Therefore, we had a potential ten-fold difference in the amount of bacterial DNA that was input to
our analysis. Initial sample centrifugation speeds and times also differed between the two groups.

This comparison raises two possible hypotheses. Firstly, amniotic fluid bacterial colonisation or accumulation might be dynamic across gestation. Secondly, it might vary across populations. To clarify the associations observed in this thesis, geographically and demographically matched mothers should be compared at mid- and full-gestation. Ideally, mothers undergoing both mid-gestation amniocentesis and full-gestation Caesarean section delivery would be sampled for this purpose (as it would not be ethically possible to obtain serial samples at other points during gestation). It should be noted that studies such as these are more difficult to perform now that non-invasive prenatal testing has reduced the amount of amniocentesis sampling performed. Further, it would be of great interest to compare the amniotic fluid microbiota of demographically matched mothers delivering by Caesarean section in different geographical locations around the world. In this manner we may develop a more comprehensive understanding of pre-natal microbial exposures.

**Limitations**

Although the definition for “microbiome” explicitly includes all microbes (i.e. viruses, bacteria, fungi, archaea, and protozoa), microbiome research has been primarily focused on bacteria. Indeed, “microbiome” has become a term that is synonymous with “bacterial microbiome”. This is because 16S rRNA gene sequencing technology is relative cheap and simple to perform. Further, there are large and well annotated databases of 16S rRNA gene sequences, allowing easy taxonomic assignment for most bacteria. The study of the non-bacterial microbiome is not as easily or widely performed. In this thesis I have described the bacterial microbiome of meconium and amniotic fluid; however, this is only one part of a larger picture. A number of recent papers have demonstrated that the non-bacterial components of the microbiome play a significant role in modulating health and disease in their human hosts [390-398]. Other studies have demonstrated that the fetus is exposed to a range of viruses, archaea, and fungi. Lim et al. identified a diverse range of bacteriophages and a limited number of eukaryotic viruses in full gestation amniotic fluid samples from healthy women [62]. This finding is particularly of note as phages are inert against humans and are, in most
cases, highly specific for their bacterial hosts. Their presence in the amniotic fluid may therefore support the theory that bacteria are also present in the amniotic fluid. Alternatively, phages may be able to pass from mother to fetus due to their small size, but may not survive or persist in the amniotic space. Wampach et al. examined colonisation and succession of bacterial, fungal, and archaeal populations in the human gut over the first year of life [13]. These authors found that meconium passed on the first day of life contained members of all three domains. Similarly, Ward et al. tracked the development of the mycobiome at multiple body sites over the first month of life, and found that all anal samples taken within a day of birth contained fungal DNA [399]. Importantly, the non-bacterial portions of the microbiome, particularly bacteriophages, might play an important role in controlling and curating the bacterial microbiome of the fetus and the intra-amniotic environment. These non-bacterial parts of the microbiome, while more challenging to study, may be transmitted to the fetus along with the bacteria described in this thesis. Although assessment of the complete microbiome was beyond the scope of this project, it is important to consider that the fetus is likely to be exposed to a broad range of microbes prior to birth, and that these might influence later-life health and disease.

A further limitation of this thesis is the use of meconium as a proxy for the fetal gut contents. Romano-Keeler et al. have previously demonstrated that the microbiota of the intestinal mucosa of neonates is distinct from matched faecal samples [160]. Unlike adults, the neonatal intestinal mucosa is actually more diverse than neonatal faeces. Therefore, the meconium microbiota described here may be an underestimation of what is truly present in the fetal gut. There are obvious practical and ethical limitations on the collection of neonatal intestinal tissue, so we have had to rely on meconium as a non-invasive measure of the fetal gut contents. Meconium is also vulnerable to contamination through feeding. To control for potential contamination from breastmilk or formula, we have only collected the first meconium passed, and only collected samples when they were passed within the first 24 h of life (mean: 8.26 h; range: 0.67 – 18.55 h). Even if feeding has commenced prior to the first meconium being passed, it is likely to only influence the microbiome of portions of meconium that sit higher up in the intestinal tract, while portions of meconium closest to the anus should remain unaffected. In support of this, we did not find any sequences belonging to Bifidobacteria sp., which have previously been identified in colostrum [400, 401], in
our meconium samples. As such, we believe that first-pass meconium is a practical, but imperfect, proxy for the fetal gut.

**Implications**

This body of work has demonstrated the importance of optimising methodologies and eliminating contamination in studies of low biomass microbiomes. This is the first perinatal microbiome study that has used dsDNase treatment of the master mix to eliminate contaminating DNA. We have shown in Chapter 4 that such treatments reduced the number of contaminating reads in our workflow by 99%. In using this method, we have been able to examine the amniotic fluid and meconium microbiomes with a higher level of sensitivity than ever before. One of the primary findings of this thesis is that the meconium microbiome does not harbour a great variety of bacteria, nor does it vary a great deal between individuals. Given that reagent-derived contamination is sporadic and often highly diverse, it is possible that previous descriptions of the meconium microbiome have actually been descriptions of the “kitome” and the “mixome”. Studies that do not include contamination controls should be interpreted with great caution.

Our findings strongly suggest that the term fetus is not sterile, while in early pregnancy it is likely that the majority of fetuses are not exposed to bacteria. Contact with bacteria and/or bacterial products prior to birth would be expected to have major effects on the developing fetal digestive tract, skin, respiratory tract, immune system, and brain. These findings have implications for our understanding of microbial colonisation and succession, and immune development in humans. By understanding the timing and nature of prenatal microbial exposures, and the influences that these have on postnatal health and disease, we may allow in-utero manipulation to optimise the foundations of the human microbiome. It seems plausible that humans are exposed to small quantities of microbes and microbial products prior to birth to prepare the developing immune systems for the massive microbial onslaught they will immediately encounter in extra-uterine life. Maternal-fetal transfer of bacteria prior to birth might be the first step of the process of immune-education that continues outside the uterus [402]. Miller described the co-dependent, symbiotic relationships that exist in the eukaryotic holobiont as an
“arc of conjoined life” [403]. The work presented here supports the theory that this arc is formed in fetal life. Aberrations to the foundation of this arc, due to maternal illness or dysbiosis, antibiotic use, or preterm birth, might influence successive postnatal colonisation and early-life immune development.

Among those who ascribe to the *in-utero* colonisation theory, there is debate around the source of the intra-amniotic microbiome. Evidence from animal models strongly suggests that the maternal gut is the primary source of bacterial seeding [64, 78, 118]. In cases of pathogenic infections of the amniotic cavity, both oral [404] and vaginal bacteria [325, 405, 406] have been implicated, demonstrating their ability to invade the amniotic space. The major bacteria detected in the studies within this thesis are not typically thought of as human gut, vaginal, or oral colonisers. *Lactobacillus*, the major genus found in the vaginal microbiome, was absent from all but one sample in this study. A single meconium sample contained 7,017 reads that mapped to *Lactobacillus iners* with a 99% sequence homology. Although we have not compared these samples with matched maternal vaginal samples, we can speculate that this particular species has ascended from the vagina and colonised this fetus. Interestingly, the presence of lactic acid bacteria in meconium has previously been associated with an increased risk of development of infant respiratory disease [24]. This may suggest that fetal exposure to vaginal lactic acid bacteria might, counterintuitively, pose a health risk, although such a finding needs to be replicated. The data presented in this thesis suggests that this type of exposure is not a common occurrence. No other meconium or amniotic fluid samples contained members of the *Lactobacillus* genus, suggesting that in most cases the vaginal microbiome does not play a role in seeding the fetus prior to birth. The bacterial profiles reported here most closely resemble those reported in the endometrial microbiome [291, 292]. Verstraelen *et al.* collected samples from the surface of the endometrium of non-pregnant women using a sheathed brush that limited the possibility of cervico-vaginal contamination of the samples [292]. These authors identified *Pelomonas sp.* in all 19 samples, and found that the genus made up around a third of all reads. Fang *et al.* took endometrial swabs from 30 women using a similar sheathed brush [291]. These authors also detected *Pelomonas sp.* in their samples, albeit in much lower abundance. However, the apparent endometrial microbiome reported in this study closely resembles the vaginal microbiome, possibly indicating contamination during sampling. Neither of these studies controlled for reagent-based contamination, so it is difficult to make
conclusions on the occurrence of Pelomonas sp. in the endometrium. Nevertheless, our data might point to the endometrium as the source of the fetal microbiome.

**Directions for future research**

*P. puraquae* was highly abundant in our meconium samples. In many cases *P. puraquae* was the only species detected in these samples. Highly abundant sequences can overwhelm PCR reactions, leading to an under-estimation of diversity. Going forward, we could design blocking primers to prevent amplification of *P. puraquae* DNA to allow less abundant sequences to be amplified. This approach has previously been successfully used to detect low abundance species in ticks and krill [407, 408]. Unfortunately, the samples analysed in this thesis have been exhausted, so such an approach cannot be used here. To validate the presence of *P. puraquae* in these samples, visual techniques such as fluorescent in-situ hybridisation might be employed. However, meconium is an opaque, insoluble, sticky, tar-like substance that does not lend itself well to such work. Faecal matter produced later in neonatal life (after feeding has commenced) is more liquescent and easier to work with. However, such samples do not represent the fetal gut contents *in-utero* and are therefore not useful for answering the questions posed in this thesis. Cultivation of *P. puraquae* from first-pass meconium samples would prove not only its presence, but also its viability. To date *P. puraquae* has only been cultured once, in the study in which it was first described.

The data presented in Chapter 6 of this thesis raise the possibility that fetal microbial exposure does not follow a universal pattern, but, rather, varies with geography. Australia is a remote and isolated island. Perth, the city in which this study was conducted, is the second most isolated capital city in the world, surrounded by ocean on one side and desert on the other. It is, therefore, unsurprising to find that the microbiome of pregnant women from Perth might vary from those located elsewhere. Previous studies have highlighted the extent to which the human microbiome varies geographically [282, 283, 285], limiting our ability to draw broad conclusions from a single population. There is, therefore, a need to perform carefully controlled microbiome analyses, such as the work described here, in a variety of populations across the globe.
For the purposes of the studies presented here, we collected samples from women delivering by elective Caesarean section only. The rationale for this approach was to allow sampling of the amniotic fluid and placenta without the risk of contamination by the vaginal microbiota. While conflicting evidence exists on this point [409], it is possible that the microbiology of the first pass meconium could be affected by delivery mode. This is likely to be a consequence of intrapartum antibiotic use [409]. For consistency and ease of recruiting, we also only enrolled women who were delivering by planned elective Caesarean section with no onset of labour. It is possible that the fetus might be rapidly colonised by vaginal microbiota following rupture of membranes. It would be interesting to compare first pass meconium samples from matched neonates born from elective Caesarean section, emergency Caesarean section, vaginal delivery with intrapartum antibiotics, and vaginal delivery without intrapartum antibiotics. This would allow the impact of antibiotics, labour, and delivery mode on the neonatal microbiome to be examined independently.

The maternal microbiome is a potential target for therapeutic manipulation during pregnancy. Animal studies have demonstrated the ability of ingested microbes to access the fetus and the amniotic fluid [64, 78, 118]. Therefore, maternal probiotic or prebiotic administration during gestation might be able to influence the fetal microbiome. Large, well-designed, randomised controlled trials of pro/prebiotic supplementation in pregnancy and its influence on the fetal or infant microbiome are currently underway. A recent meta-analysis revealed that antenatal pro/prebiotic supplementation has no discernible impact on adverse pregnancy outcomes [410]. A simplistic view of the potential use of pro/prebiotics in pregnancy should not be taken. New research suggests that the human gut is resistant to probiotics, with commercial formulations limited in their ability to colonise a healthy host [411]. Simple commercial formulations of probiotics are, therefore, unlikely to have much effect on the fetal microbiome. A detailed understanding of the mechanisms behind fetal microbiome seeding is required to take advantage of this early-life window of opportunity.

Bacteria present in the amniotic fluid are likely to be the source of initial seeding for the skin microbiome. The OTUs detected in amniotic fluid in this thesis largely mapped to bacteria that are found in the adult skin microbiome. However, *P. puraquae*, the major
OTU detected in meconium samples, is not a major component of the infant or adult gut microbiome. The data presented in Chapter 5 capture the fetal and intra-amniotic microbiome at a single moment in time – birth. Going forward, it would be interesting to track succession patterns in this population. The human gut undergoes rapid colonisation following birth [100, 372, 402]. Bacteria present in the gut at birth are likely to be displaced by secondary postnatal colonisers as the contents of the infant gut changes. Alternatively, founding species may remain present in small quantities, but be outnumbered by postnatal colonisers. The extent to which pioneering species of the gastrointestinal tract influence subsequent colonisation events is a question that remains under-investigated. Two recent studies have used strain-level resolution metagenomics to track colonisation and succession patterns in mother-infant pairs. Yassour et al. collected meconium samples on the first day of life from 44 infants [61]. They found these samples to be of very low complexity, usually dominated by a single phylum and only 3-7 species. They also reported that the meconium microbiome was strikingly different to the maternal gut microbiome, but that it nevertheless contained a number of shared species. These authors then followed the infant and maternal gut microbiomes over three months and observed two patterns of postnatal maternal transmission of bacteria. In the first, the dominant strains present in the mother were passed to the infant. In the second, less dominant strains that conferred a functional benefit were passed to the infant. These data suggest that initial colonisation is driven by stochastic processes, while secondary colonisation is driven by functional selection. Ferretti et al. also sampled meconium on the first day of life in 25 neonates [60]. These authors reported that the meconium microbiome contained a high level of inter-subject variability, but that the majority of strains detected were shared with at least one maternal body site (skin, breast milk, vagina, gut, or oral cavity), with the maternal gut being the largest donor. In the four months following birth, inter-subject variability decreased, and then gradually increased again. Again, these findings point to a stochastic pattern of initial colonisation that becomes deterministic in the first months of life. Importantly, these authors found that some species that existed in meconium in the first day of life were lost in subsequent samples, while others persisted over time. Similarly, Wampach et al. found that meconium samples were richer and more diverse than infant stool samples, and that richness and diversity decreased after birth, before gradually increasing again [13]. These authors also found that some species that were present in meconium samples on day one of life were lost shortly thereafter, while
others persisted throughout the first year. Together, these data demonstrate the need for high-resolution longitudinal studies to examine microbial succession in the infant gut.

**The not-so-sterile womb: new data to challenge an old dogma**

The theory that the womb and the fetus are sterile has been hotly debated over the past decade. The tide of evidence is now turning in favour of the *in-utero* colonisation hypothesis. However, questions relating to the validity of microbiome data produced from low biomass samples still hang heavy in this field. In this thesis the fetal microbiome has been characterised as never before. Here I have presented the first full-length 16S rRNA gene sequencing survey of the fetal gut and amniotic fluid, and association with maternal health parameters and fetal immune status. By taking the time to optimise the methods used here, and to eliminate contaminating DNA, I have been able to create a sensitive and accurate portrait of the birth of the holobiont. The data presented in this thesis lays the groundwork for future hypothesis-driven research in this field. Importantly, this work has exciting implications for our own definitions of self, by demonstrating that the scaffolding of the microbiome that we will carry with us for life develops prior to birth.
Reference list


182


220. Mirsepasi, H., S. Persson, C. Struve, L.O. Andersen, A.M. Petersen, and K.A. Krogfelt, Microbial diversity in fecal samples depends on DNA extraction method: Easymag DNA extraction compared to qiaamp DNA stool mini kit extraction. BMC Res Notes, 2014. 7: p. 50.


196


Appendix I: GHAP code

Greenfield Hybrid Amplicon Pipeline (GHAP) - v2.1 - 15/4/2018 - Windows

Paul Greenfield CSIRO (paul.greenfield@csiro.au)

rem Run options
rem *******
rem *** type of amplicons - 16S, 18S or ITS, 18S-PR2
rem if you have a custom reference set, give it a name here and define
its files below
SET type=16S
rem *** name used in creation of results files (type will be added XXX
--> XXX_type)
SET OTU=FMS
rem full run (no), or stop after read lengths histogram generated
(yes) to allow for trimPM to be set properly.
SET stopAfterHisto=no
rem should the data be demultiplexed? no=already done;
Illumina=standard Illumina; custom=custom barcodes with possible
leading Ns.
SET dmux=no
rem SET dmux=Illumina
rem SET dmux=custom
rem merge read pairs? Only set to 'no' if merging has already been
done, or not relevant (such as 454)
SET mergeReads=no
rem and keep the un-merged but good-looking R1 reads - might be needed
for ITS if the insert cannot be covered by a pair of reads
SET keepR1=no
rem should already-merged or unpaired reads be filtered? Not used with
pair-merging.
SET filterReads=no
rem filter OTUs before classification to remove non-target-gene
sequences
SET filterOTUs=yes
rem classify reads with RDP (as well as OTUs)?
SET classifyReads=no
rem denoising parameters for cut-down OTU tables (no thinned files are
produced if this is no). Discards small and unreliable OTUs from the
tail
REM SET denoise=no
SET denoise=DN_1-3
SET denoiseParams=--minPerSample 3 --minSamples 1
rem clustering or zOTUs?
SET folding=cluster
REM SET folding=zOTU

rem Installation options
rem *******
rem Directory where GHAP code and data files were installed
SET code=C:\Work\GHAPv2.1\GHAP_code
rem # code=$HOME/GHAPv2
rem RDP classifier (used for 16S and ITS)
SET classifier=%code%\rdp_classifier_2.12\dist\classifier.jar
rem Usearch location (see below for Usearch compatibility options))
SET usearch=%code%\usearch10.0.240_win32.exe
rem Usearch options: V10 broke compatibility with earlier releases.
Choose the parameter set here that matches the version of USearch.
rem <v10
REM SET mergeParams=--fastq_minovlen 25 --fastq_maxdiffpct 20
REM SET derep=derep_fulllength
REM >=v10
SET mergeParams=-fastq_minovlen 25 -fastq_pctid 80
SET derep=fastx_uniques
rem how many parallel threads to use (for ParallelForFiles and USearch). Adjust to reflect the number of cores available.
SET OMP_NUM_THREADS=16
rem Don't change the following lines
rem ensure the code directory is included in the search path
SET path=%path%;%code%
rem # PATH=$PATH:$code
SET threads=%OMP_NUM_THREADS%
rem partition size (used only for very large datasets such as full HiSeq lanes).
SET partSize=2000000
rem

rem Various directories and file names/bits
rem ------------------------------------------------------
rem *** MUST set temp directory - fastest available storage is best. May NOT be the same directory as dataDir
SET temp=C:\Users\gre403\Documents\GHAPTemp\%OTU%
rem # temp=$tmp\[412]
rem remember where we are so we can go back here at the end
SET startingDir="%CD%"
rem *** directory where the data files are to be found (if not the current directory)
SET dataDir=%startingDir%
rem *** pattern to find all reads files - expects R1 and R2 naming
SET reads=%dataDir%*_*_.fastq
rem *** R1 reads mark (typically R1_* or R1 - should go all the way to the .fastq/...)
SET R1=R1_*
REM SET R1=R1
rem additional read file name bits to be deleted to get to sample names/directories. The specified strings will be deleted from the sample names if they're found there.
SET fnBits="-fn ccs. -fn mapped_samples_derep_t_merged_dm_ -fn _noPhiX"

rem demultiplexing parameters - leave untouched if data already demultiplexed
rem ------------------------------------------------------
rem *** pattern for Illumina index files
SET index=%dataDir%Undetermined_S0_L001_I1_001.fastq
rem *** pattern for mapping file (pattern allowed to return multiple files, each pair of read files will then be matched to the next mapping file)
SET mapping=%dataDir%\mapping_file.txt
rem *** col# of sampleId in mapping file (Illumina & custom barcodes)
SET idCol=0
rem *** col# of forward barcode in mapping file (Illumina multiplexing)
SET bcCol=1
rem *** col# of forward barcode in mapping file (custom barcodes)
SET fbcCol=1
rem *** col# of forward primer in mapping file (custom barcodes)
SET fpcCol=2
rem *** col# of reverse barcode in mapping file (custom barcodes)
SET rbcCol=3
rem *** col# of reverse primer in mapping file (custom barcodes)
SET rpcCol=4
rem Trimming parameters
rem -------------------
rem *** optional start trimming for all files (trims N bases from
start of each read)
SET trimStart=0
rem *** optional qual-driven trimming for end of reads (run fastqc for
cutoff; 1=no qual trimming)
SET trimQual=1
rem *** minimum/maximum length for merged reads
rem nnn says to drop any reads shorter than nnn bases, and truncate
any longer ones to this length
rem nnn+ says to keep any reads at least this long, reads shorter
than this are dropped
rem nnn-mmm says to keep any reads between nnn and mmm bases
long, and to drop reads outside this length range
rem nnn++mmm says to keep any reads greater than nnn bases, and
to trim any longer than mmm
SET trimPM=1400-1600
rem Only used if un-merged R1 reads are being folded back into merged
reads. Fixed R1 reads length after trimming (choose from histogram).
SET trimR1=240

rem Per-data type files/parameters
rem -------------------------------
rem You can add your own classification parameters here. You shouldn't
need to touch these settings unless you want to customise the
classification part of the pipeline
SET class=none
SET classifyDB=none
SET blastDB=none
SET conf=0.5
SET evalue=1E-5

rem PhiX filter file name
SET PhiXFilter=%code\PhiX_20.mer
SET PhiXBLASTdb=%code\PhiX_BLAST.udb

if %type% equ 16S (  
    SET class=classify  
    SET RDPtrs=16srRNA  
    SET classifyDB=%code\P16S\RDPv16+RefSeq_5-18_16S.udb  
    SET geneFilter=%code\P16S\RDPv16+RefSeq_5-18_16S_NC_20.mer  
    SET conf=0.5  
)

if %type% equ ITS (  
    SET class=classify  
    SET RDPtrs=fungalits_warcup  
    SET classifyDB=%code\WITS\Warcup_fungal_ITS_V2.udb  
    SET geneFilter=%code\WITS\Warcup_fungal_ITS_V2_20.mer  
    SET conf=0.5  
)

if %type% equ 18S (  
    SET class=blast  
    SET blastDB=%code\SILVA18S\SILVA_128_18S_Cv2.0_BLAST.id.udb  
    SET blastTaxo=%code\SILVA18S\SILVA_128_18S_Cv2.0_taxonomy.txt  
    SET geneFilter=%code\SILVA18S\SILVA_128_18S_Cv2.0_20.mer  
)
SET geneBLASTdb=%code%\SILVA18S\SILVA_128_18S_Cv2.0_BLAST_species.udb
SET evalue=1E-5
SET BLASTQueryCov=0.8
SET taxoBLASTCutoffs="-sk 1.0 -k 1.0 -p 77.0 -sp 77.0 -o 85.0 -f 90.0 -g 95.0 -s 97.0"
)

if %type% equ 18S-PR2 ( SET class=blast
SET blastDB=%code%\PR2\pr2_4.6_Eukaryota_BLAST_id.udb
SET blastTaxo=%code%\PR2\pr2_4.6_Eukaryota_taxonomy.txt
SET geneFilter=%code%\PR2\pr2_4.6_Eukaryota_20.mer
SET geneBLASTdb=%code%\PR2\pr2_4.6_Eukaryota_BLAST_species.udb
SET evalue=1E-5
SET BLASTQueryCov=0.8
SET taxoBLASTCutoffs="-sk 1.0 -k 1.0 -p 77.0 -sp 77.0 -c 80.0 -o 85.0 -f 90.0 -g 95.0 -s 97.0"
)

if %type% equ COI ( SET class=blast
SET blastDB=%code%\COI\GenBank_1-5-18_COI_id_C99_BLAST.udb
SET blastTaxo=%code%\COI\GenBank_1-5-18_COI_taxonomy_derep_C99.txt
SET geneFilter=%code%\COI\GenBank_1-5-18_COI_taxonomy_derep_C99_20.mer
SET geneBLASTdb=%code%\COI\GenBank_1-5-18_COI_species_C99_BLAST.udb
SET evalue=1E-5
SET BLASTQueryCov=0.8
SET taxoBLASTCutoffs="-sk 1.0 -k 1.0 -p 77.0 -sp 77.0 -c 80.0 -o 85.0 -f 90.0 -g 95.0 -s 97.0"
)

if %class% equ none ( echo unknown data type %type% exit /b )

rem ----------- start of clustering/classification script -----------

rem try to create the temp directory (in case it doesn't exist already) if not exist %temp% ( mkdir %temp% )
rem change to the Temp directory cd /d %temp%

if %dmux% equ Illumina ( echo --- demultiplexing reads in standard Illumina format rem The data is expected to be in normal Illumina format with an index file and R1 & R2 reads files rem You may have to adjust the idCol and bcCol to match your mapping file (zero-based indexes for id and barcode columns) rem This will demultiplex the reads (and remove PhiX) - and leave the demultiplexed files in Temp DemultiplexMiSeq -m "%mapping%" -i "%index%" -f %PhiXFilter% -o %temp% -id %idCol% -bc %bcCol% "%reads%"
rem remove the reads that couldn't be assigned to a barcode del *_notFound.*
)
if %dmux% equ custom (}
echo --- demultiplexing using custom variable-length barcodes
DemultiplexCustomBC -m "%mapping%" -o %temp% -n %type% -id %idCol% -fbc %fbcCol% -fpc %fpcCol% -rbc %rbcCol% -rpc %rpcCol% "%reads%"
)
if %dmux% equ no (rem already demultiplexed so just copy the files to Temp
echo --- reads already demultiplexed, copying to Temp
ParallelForFiles -t 4 "%reads%" CatSeqs @file dm_@fname@ext
)
rem merging paired reads (needs v8.1+ of Usearch)
if %mergeReads% equ yes (echo --- trimming reads before merging
rem trim away the start of each read if these look like they
could be adapters etc; and any bad bases at the end (poor qual scores)
ParallelForFiles -t 4 dm_*.* TrimReads @file +%trimStart% -qual %trimQual% t_@fname.fastq
del dm_*.*
)
if %keepR1% equ no (echo --- merging read pairs
rem merge paired FASTQ (overlapping reads) and convert
merged reads to FASTA
ParallelForFiles -t %OMP_NUM_THREADS% t_dm_*%R1%.fastq %usearch% -fastq_mergepairs @file -fastaout merged_@fname.fa %mergeParams%
del t_dm_*.*fastq
)
if %keepR1% equ yes (echo --- merging read pairs and keeping good un-merged R1
reads
rem merge paired FASTQ (overlapping reads), convert merged
reads to FASTA and save the R1 reads that could not be merged
successfully
ParallelForFiles -t %OMP_NUM_THREADS% t_dm_*%R1%.fastq %usearch% -fastq_mergepairs @file -fastaout merged_@fname.fa -fastqout_notmerged_fwd notmerged_R1_@fname.fastq %mergeParams%
rem filter out the worst quality reads from the R1
'discards'
ParallelForFiles -t %OMP_NUM_THREADS% t_dm_*%R1%.fastq %usearch% -fastq_filter notmerged_R1_@fname.fastq -fastq_maxee 3.0 -fastaout filtered_notmerged_R1_@fname.fa
rem only keep reads that don't appear to be shortened
versions of one we've already merged
ParallelForFiles -t %OMP_NUM_THREADS% t_dm_*%R1%.fastq %usearch% -usearch_global filtered_notmerged_R1_@fname.fa -db merged_@fname.fa -strand both -id 0.95 -maxaccepts 64 -maxrejects 256 -notmatched notmatched_filtered_notmerged_R1_@fname.fa
rem finally filter them and only keep those that look like
they're from the target organisms/region
ParallelForFiles -t %OMP_NUM_THREADS% t_dm_*%R1%.fastq %usearch% -ublast notmatched_filtered_notmerged_R1_@fname.fa -db %geneBLASTdb% -strand both -evalue 1e-6 -query_cov 0.8 -maxaccepts 64 -maxrejects 512 -top_hit_only -matched matched_notmatched_filtered_notmerged_R1_@fname.fa
rem generate a histogram of the R1 reads so that the full
run can trim them correctly
rem trim R1 reads to the same length to avoid problems with clustering (same seq but different length will go into different clusters)
rem these filtered, trimmed R1 reads will be added to the merged reads before clustering
ParallelForFiles -t 4 t_dm_*.%R1%.fastq
TrimReadsToFixedLength
matched_notmatched_filtered_notmerged_R1_@fname.fa %trimR1%
t_matched_notmatched_filtered_notmerged_R1_@fname.fa

rem reads already merged or not paired at all.
if %mergeReads% equ no ( 
  echo --- trimming already-merged reads
  rem quality filter and then trim
  if %filterReads% equ yes ( 
    rem trim first in case we're getting rid of error-prone regions
    ParallelForFiles -t 4 dm_*.* TrimReads @file +%trimStart% -qual %trimQual% t_@fname.fastq
    rem and remove phiX if present
    ParallelForFiles -t 4 t_dm_*.* FilterReads -r noPhiX -f
    %PhiXFilter% 25pct @file
    del dm_*.*
    rem and then qual filter
    ParallelForFiles -t %OMP_NUM_THREADS% t_*_noPhiX.fastq %usearch% -fastq_filter @file -fastq_maxee 5.0 -fastaout
    merged_@fname.fa
    del t_*_.fastq
  )
  else ( 
    rem or just trim away the start of each read (they could be adapters etc), and convert to reads to FASTA if necessary
    ParallelForFiles -t 4 dm_*.* FilterReads -r noPhiX -f
    %PhiXFilter% 25pct @file
    ParallelForFiles -t 4 dm_*_noPhiX.* TrimReads -fasta @file +%trimStart% -qual %trimQual% merged_@fname.fa
del dm_*.*
  )
)

rem Use this histogram to choose a good value for trimPM - the length of the trimmed reads going into the clustering steps
HistoReadLengths matched_notmatched_filtered_notmerged_R1_*.fa
%OTU%_%type%_R1_notmerged_LengthsHisto.txt
  copy %OTU%_%type%_R1_notmerged_LengthsHisto.txt
%HistoReadLengths
matched_notmatched_filtered_notmerged_R1_*.fa
%OTU%_%type%_LengthsHisto.txt
copy %OTU%_%type%_LengthsHisto.txt %startingDir%
del %OTU%_%type%_LengthsHisto.txt

rem exit now if we need to look at the histogram and set trimPM correctly
if %stopAfterHisto% equ yes ( 
  cd /d %startingDir%
  exit /b
)
rem discard any short reads and trim longer reads if requested
echo --- trimming merged reads
ParallelForFiles -t 4 merged_*\.fa TrimReadsToFixedLength @file
%trimPM% t @fname\.fa
del merged_*\.fa

rem add in the already-trimmed etc R1 reads
if %keepR1% equ yes (    
rem and add the R1 reads to the merged reads for the clustering
steps
    ParallelForFiles -t 2 t_dm_*\.R1%\fastq CatSeqs -append
t_matched notmatched_filtered_notmerged_R1 @fname\.fa
t_merged @fname\.fa
del t_dm_*\.fastq
)
echo --- partitioning and dereplicating reads
rem an initial derep of each sample to reduce the size of each per-
rem sample file
rem these derepped trimmed reads are the ones that will be used later
when reporting on read numbers
ParallelForFiles -t %OMP_NUM_THREADS% t_merged_*\.fa %usearch% -%derep%
@file -fastaout samples_derep_@fname\.fa -sizeout
del t_merged_*\.fa

rem combine all the trimmed, merged files into a single file
CatSeqs samples_derep_t_merged_*\.fa combined_t_%OTU%\.fa

rem partition all of the trimmed, merged, filtered reads into big
rem chunks for OTU clustering steps
PartitionAmplicons combined_t_%OTU%\.fa %partSize% p
del combined_t_%OTU%\.fa

rem dereplicate each partition (although there may only be one)
ParallelForFiles -t 2 p_*\.combined_t_%OTU%\.fa %usearch% -%derep% @file
-threads %threads% -fastaout derep_@fname\.fa -sizein -sizeout
del p_*\.combined_t_%OTU%\.fa

rem if we have multiple partitions, go through the derep cycle again
to remove common seqs across partitions
if exist derep_p_1_combined_t_%OTU%\.fa (    
rem combine, repartition and derep the dereplicated reads once
again
    echo repeating partitioning and dereplication
rem put sequences with similar rep counts together
RepartitionAmplicons derep_p_*\.combined_t_%OTU%\.fa ps
del derep_p_*\.combined_t_%OTU%\.fa
ParallelForFiles -t 2 ps_derep_p_*\.combined_t_%OTU%\.fa %usearch%
-%derep% @file -threads %threads% -fastaout derep_@fname\.fa -sizein -sizeout
del ps_derep_p_*\.combined_t_%OTU%\.fa
rem one more partition and derep
CatSeqs derep_ps_derep_*\.fa combined_t_%OTU%\.fa
del derep_*\.fa
PartitionAmplicons combined_t_%OTU%\.fa %partSize% p
del combined_t_%OTU%\.fa
ParallelForFiles -t 2 p_*\.combined_t_%OTU%\.fa %usearch% -%derep%
@file -threads %threads% -fastaout derep_@fname\.fa -sizein -sizeout
del p_*\.combined_t_%OTU%\.fa
)

rem if we still have multiple partitions, cluster and derep and
pretend we have a single derepped 'reads' file at the end
if exist derep_p_1_combined_t_%OTU%\.fa (
echo --- derep/clustering partitions to further reduce size

rem sort each partition prior to clustering (minSize=2 here will drop a small number of rare sequences but we're only here because we have a very large number of seqs in total)

ParallelForFiles -t 2 derep_p_*_combined_t_%OTU%.fa %usearch% -threads %threads% -sortbysize @file -fastaout sorted@fname.fa -minsize 1
del derep_p_*_combined_t_%OTU%.fa

rem cluster each sorted, derepped partition
if %folding% equ cluster
    ParallelForFiles -t %OMP_NUM_THREADS% sorted_derep_p_*_combined_t_%OTU%.fa %usearch% -cluster_otus @file -otus clusters@fname.fa -relabel OTU@fname_
if %folding% equ zOTU
    ParallelForFiles -t %OMP_NUM_THREADS% sorted_derep_p_*_combined_t_%OTU%.fa %usearch% -unoise3 @file -zotus clusters@fname.fa
)

rem combine the OTU partitions
echo --- combine clustered partitions and re-derep again
CatSeqs clusters_sorted_derep_p_*_combined_t_%OTU%.fa
combined_clusters%OTU%_%type%_OTUs.fa
del clusters_sorted_derep_p_*_combined_t_%OTU%.fa

rem derep the combined cluster seqs to fold together the same seqs from multiple partitions
%usearch% -derep combined_clusters%OTU%_%type%_OTUs.fa -threads %threads% -fastaout derep_combined_clusters%OTU%_%type%_OTUs.fa
del combined_clusters%OTU%_%type%_OTUs.fa

rem map the partitioned derepped reads back onto this combined, derepped, clustered set of seqs in preparation for getting accurate read counts
ParallelForFiles -t 1 sorted_derep_p_*_combined_t_%OTU%.fa
%usearch% -usearch_global @file -threads %threads% -db derep_combined_clusters%OTU%_%type%_OTUs.fa -threads %threads% -strand plus -id 0.97 -uc mapped@fname.uc -maxaccepts 64 -maxrejects 256 -top_hit_only

rem and use these mapping files (.uc) to add size annotations to the derepped, combined cluster seqs - and create a new derep_p_0_combined_t_%OTU%.fa

rem --- create final set of sized, derep reads for final clustering
CountReadsForOTUs derep_combined_clusters%OTU%_%type%_OTUs.fa
derep_p_0_combined_t_%OTU%.fa
mapped_sorted_derep_p_*_combined_t_%OTU%.uc
del sorted_derep_p_*_combined_t_%OTU%.fa
del mapped_sorted_derep_p_*_combined_t_%OTU%.uc
del derep_combined_clusters%OTU%_%type%_OTUs.fa

rem finally size-sort the dereplicated reads prior to clustering step
echo --- sort prior to clustering and discard singletons
%usearch% -sortbysize derep_p_0_combined_t_%OTU%.fa -fastaout sorted_derep_p_0_combined_t_%OTU%.fa -minsize 1
del derep_p_0_combined_t_%OTU%.fa
rem cluster the reads/centroids together to produce the final set of OTUs
echo cluster to get final set of OTUs
if %folding% equ cluster (  
  %usearch% -cluster_otus sorted_derep_p_0_combined_t_%OTU%.fa -otus lb_%OTU%_%type%_OTUs.fa -relabel OTU -minsize 1  
)
if %folding% equ zOTU (  
  %usearch% -unoise3 sorted_derep_p_0_combined_t_%OTU%.fa -zotus lb_%OTU%_%type%_OTUs.fa  
)
del sorted_derep_p_0_combined_t_%OTU%.fa

echo --- map merged reads back onto OTUs and correct OTU headers
rem Map merged, trimmed reads (including singletons) back to the final OTUs to get final counts of reads assigned to each OTU
rem Use '-strand plus' as this is faster and gives the same results - due to nature of amplicon PCR and merging R2 onto R1 reads.
echo mapping reads back to OTUs to get counts
ParallelForFiles -t 2 samples_derep_t_merged_*.fa %usearch% -usearch_global @file -db lb_%OTU%_%type%_OTUs.fa -threads %threads% -strand plus -id 0.97 -uc mapped_f@fname.uc -maxaccepts 64 -maxrejects 256 -top_hit_only

rem Update the read counts in the OTU headers
echo updating read counts for OTUs
CountReadsForOTUs lb_%OTU%_%type%_OTUs.fa sized_lb_%OTU%_%type%_OTUs.fa mapped_samples_derep_t_merged_*.uc
del lb_%OTU%_%type%_OTUs.fa
rem And sort OTUs by size again
echo sorting OTUs by size
%usearch% -sortbysize sized_lb_%OTU%_%type%_OTUs.fa -fastaout %OTU%_%type%_OTUs.fa
del sized_lb_%OTU%_%type%_OTUs.fa

rem filter the OTUs if requested (and keep the discards in case someone wants to look at them)
if %filterOTUs% equ yes (  
  echo filtering reads
  FilterReads -r only -full -discards +fz %geneFilter% 20pct -f %PhiXFilter% 50pct %OTU%_%type%_OTUs.fa  
  copy /Y %OTU%_%type%_OTUs_only.fa %OTU%_%type%_only_OTUs.fa  
  del %OTU%_%type%_OTUs_only.fa  
  if exist %OTU%_%type%_OTUs_discards.fa (  
    copy /Y %OTU%_%type%_OTUs_discards.fa %OTU%_%type%_OTUs_discards_OTUs.fa  
  )  
  del %OTU%_%type%_OTUs_discards_OTUs.fa  
)

if %class% equ classify (  
  echo --- classify %type% OTUs and generate OTU table
  %usearch% -ublast %OTU%_%type%_OTUs_discards_OTUs.fa -db %geneBLASTdb% -evalue %evalue% -alnout %OTU%_%type%_OTUs_discards_OTUs.ali -uc %OTU%_%type%_OTUs_discards_OTUs.uc -strand both -top_hit_only -maxaccepts 64 -maxrejects 256 -query_cov 0.5  
  %usearch% -ublast %OTU%_%type%_OTUs_discards PhiX_OTUs.ali -uc %OTU%_%type%_OTUs_discards PhiX_OTUs.ali -strand both -top_hit_only -maxaccepts 64 -maxrejects 256 -query_cov 0.5  
)
rem Classify the OTUs with RDP
ClassifyOTUs -min %conf% -p 1000 -t %OMP_NUM_THREADS% -c %classifier% -g %RDpts% %OTU%_%type%_OTUs.fa OTUs_%OTU%_%type%.ass
OTUs_%OTU%_%type%_classified_%conf%.txt
echo --- matching OTUs to reference seqs matching OTUs to reference seqs
%usearch% -usearch_global %OTU%_%type%_OTUs.fa -db %classifyDB%
-threads %threads% -strand both -id 0.85 -uc OTUs_%OTU%_%type%_species_RDP.uc -maxaccepts 256 -maxrejects 256 -top_hits_only

rem Generate full OTU table
GenerateOTUTable -s OTUs_%OTU%_%type%_species_RDP.uc -c OTUs_%OTU%_%type%_classified_%conf%.txt
mapped_samples_derep_t_merged_*.uc %OTU%_%type%_OTUTable.txt -fn mapped_samples_derep_t_merged_t_dm_ -fn .uc %fnBits%

rem fill empty spots in OTU taxonomy part with _XXX
FillRDPOTUTable %OTU%_%type%_OTUTable.txt
%OTU%_%type%_OTUTable_usX.txt

rem Convert OTU tables to .biom format
ConvertTableToBIOM %OTU%_%type%_OTUTable.txt
%OTU%_%type%_OTUTable_usX.biom

rem Summarise the OTU table at various taxonomic levels
SummariseRDPOTUTable %OTU%_%type%_OTUTable.txt
%OTU%_%type%_OTUTable_summary.txt
SummariseRDPOTUTable %OTU%_%type%_OTUTable_usX.txt
%OTU%_%type%_OTUTable_usX_summary.txt

rem generate results that only cover the OTUs that look like they belong to the target gene (filtering)
if %filterOTUs% eq yes ( ThinOTUTable -fa %OTU%_%type%_only_OTUs.fa %OTU%_%type%_OTUTable.txt %OTU%_%type%_only_OTUTable.txt
ConvertTableToBIOM %OTU%_%type%_only_OTUTable.txt
%OTU%_%type%_only_OTUTable_usX.txt
ConvertTableToBIOM %OTU%_%type%_only_OTUTable_usX.txt
SummariseRDPOTUTable %OTU%_%type%_only_OTUTable.txt
%OTU%_%type%_only_OTUTable_summary.txt
FillRDPOTUTable %OTU%_%type%_only_OTUTable.txt
%OTU%_%type%_only_OTUTable_usX.txt
ConvertTableToBIOM %OTU%_%type%_only_OTUTable_usX.txt
%OTU%_%type%_only_OTUTable_usX.biom
SummariseRDPOTUTable %OTU%_%type%_only_OTUTable_usX.txt
%OTU%_%type%_only_OTUTable_usX_summary.txt)

if %denoise% neq no ( rem Generate an OTU table that doesn't include the smallest OTUs
DenoiseOTUTable %denoiseParams% %OTU%_%type%_OTUTable.txt %OTU%_%type%_OTUTable %denoise%.txt
ThinOTUs %OTU%_%type%_OTUs.fa %OTU%_%type%_OTUTable %denoise%.txt %OTU%_%type%_OTUTable %denoise%.fa
ConvertTableToBIOM %OTU%_%type%_OTUTable %denoise%.biom SummariseRDPOTUTable %OTU%_%type%_OTUTable %denoise%.txt %OTU%_%type%_OTUTable %denoise%.summary.txt
FillRDPOTUTable %OTU%_%type%_OTUTable %denoise%.txt %OTU%_%type%_OTUTable %denoise%.usX.txt

211
if %filterOTUs% equ yes (  
  DenoiseOTUTable %denoiseParams%  
  %OTU%_%type%_only_OTUTable.txt  
  %OTU%_%type%_only_OTUTable_denoise%.txt  
  ThinOTUs %OTU%_%type%_OTUs.fa  
  %OTU%_%type%_only_OTUTable_denoise%.fa  
  ConvertTableToBIOM  
  %OTU%_%type%_only_OTUTable_denoise%.txt  
  %OTU%_%type%_only_OTUTable_denoise%.biom  
  SummariseRDPOTUTable  
  %OTU%_%type%_only_OTUTable_denoise%.txt  
  %OTU%_%type%_only_OTUTable_denoise%_summary.txt  
  FillRDPOTUTable  
  %OTU%_%type%_only_OTUTable_denoise%.txt  
  %OTU%_%type%_only_OTUTable_denoise%_usX.txt  
  ConvertTableToBIOM  
  %OTU%_%type%_only_OTUTable_denoise%_usX.txt  
  %OTU%_%type%_only_OTUTable%_denoise%_usX_summary.txt  
  if %classifyReads% equ yes (  
    rem classify merged,trimmed reads with RDP  
    echo --- classify reads with RDP  
    ParallelForFiles -t 2 t_*.fa ClassifyReads -min %conf% -p 500 -t %OMP_NUM_THREADS% -otus %OTU%_%type%_OTUs.ass -otumap mapped_%fname.uc -sp %OTU%_%type%_species_RDP.uc -c %classifier% -g %RDPts% -a %fname.ass %fname_RDP_%conf%.txt @file  
    rem summarise the classifications  
    MergeRDPClassifications t_*_RDP_%conf%.txt  
    %OTU%_%type%_RDP_%conf%.txt -fn t_merged_t_dm_ -fn _RDP_0.5.txt  
    %fnBits%  
  )  
  echo copy results files  
  copy %OTU%_%type%_OTUs.fa %startingDir%  
  copy %OTU%_%type%_OTUTable.txt %startingDir%  
  copy %OTU%_%type%_OTUTable.biom %startingDir%  
  copy %OTU%_%type%_OTUTable_usX.txt %startingDir%  
  copy %OTU%_%type%_OTUTable_usX.biom %startingDir%  
  copy %OTU%_%type%_OTUTable_summary.txt %startingDir%  
  copy %OTU%_%type%_OTUTable%_denoise%_usX_summary.txt %startingDir%  
  if %filterOTUs% equ yes (  
    if not exist %startingDir%\%type%_only (  
      mkdir %startingDir%\%type%_only  
    )  
    copy %OTU%_%type%_only_OTUs.fa %startingDir%\%type%_only  
    copy %OTU%_%type%_only_OTUTable.txt %startingDir%\%type%_only  
    copy %OTU%_%type%_only_OTUTable.biom %startingDir%\%type%_only  
  )

212
copy %OTU%_%type%_only_OTUTable_usX.txt
%startingDir%\type%_only
  copy %OTU%_%type%_only_OTUTable_usX.biom
%startingDir%\type%_only
  copy %OTU%_%type%_only_OTUTable_summary.txt
%startingDir%\type%_only
  copy %OTU%_%type%_only_OTUTable_usX_summary.txt
%startingDir%\type%_only
  if exist %OTU%_%type%_discards_OTUs.fa
    copy %OTU%_%type%_discards_OTUs.fa
%startingDir%\type%_only
    copy %OTU%_%type%_discards_OTUs.aln
%startingDir%\type%_only
    copy %OTU%_%type%_discards_PhiX_OTUs.aln
%startingDir%\type%_only
    copy %OTU%_%type%_discards_OTUs.uc
%startingDir%\type%_only
    copy %OTU%_%type%_discards_PhiX_OTUs.uc
%startingDir%\type%_only
  )
  if %denoise% neq no (
    if not exist %startingDir%\denoise% ( mkdir %startingDir%\denoise% )
    copy %OTU%_%type%_OTUs_%denoise%.fa
%startingDir%\denoise%
    copy %OTU%_%type%_OTUTable_%denoise%.txt
%startingDir%\denoise%
    copy %OTU%_%type%_OTUTable_%denoise%.biom
%startingDir%\denoise%
    copy %OTU%_%type%_OTUTable_%denoise%_usX.txt
%startingDir%\denoise%
    copy %OTU%_%type%_OTUTable_%denoise%_usX.biom
%startingDir%\denoise%
    copy %OTU%_%type%_OTUTable_%denoise%_summary.txt
%startingDir%\denoise%
    copy %OTU%_%type%_OTUTable_%denoise%_usX_summary.txt
%startingDir%\denoise%
    if %filterOTUs% equ yes ( copy %OTU%_%type%_only_OTUs_%denoise%.fa
%startingDir%\denoise%
      copy %OTU%_%type%_only_OTUTable_%denoise%.txt
%startingDir%\denoise%
      copy %OTU%_%type%_only_OTUTable_%denoise%.biom
%startingDir%\denoise%
      copy %OTU%_%type%_only_OTUTable_%denoise%_usX.txt
%startingDir%\denoise%
      copy %OTU%_%type%_only_OTUTable_%denoise%_usX.biom
%startingDir%\denoise%
      copy %OTU%_%type%_only_OTUTable_%denoise%_summary.txt
%startingDir%\denoise%
      copy %OTU%_%type%_only_OTUTable_%denoise%_usX_summary.txt
%startingDir%\denoise%
    )
    if %classifyReads% equ yes ( copy %OTU%_%type%_RDP_%conf%.txt %startingDir% )
  )
if exist *.ass {
```python
del *.ass

if exist *.hier
    del *.hier

if %class% equ blast
    rem Assign species id to OTUs using best BLAST match to database
    echo --- classify %type% OTUs and generate OTU table
    %usearch% -ublast %OTU%_%type%_OTUs.fa -db %blastDB% -threads %threads% -strand both -evalue %evalue% -uc %OTU%_%type%_OTUs.uc -alnout %OTU%_%type%_OTUs.aln -query_cov %BLASTQueryCov% -top_hits_only -maxaccepts 64 -maxrejects 256

    rem Generate OTU table (with speciesId)
    GenerateOTUTable -s %OTU%_%type%_OTUs.uc "mapped_samples_derep_t_merged.*.uc" %OTU%_%type%_OTUTable_ids.txt -fn mapped_samples_derep_t_merged_t_dm_-fn.uc %fnBits%
    rem Add the taxonomy info to the OTU table
    AddTaxonomyToBLASTOTUTable %OTU%_%type%_OTUTable_ids.txt %blastTaxo% %OTU%_%type%_OTUTable.txt
    del %OTU%_%type%_OTUTable_ids.txt

    rem Reduce the taxonomy using defined confidence levels to remove unreliable nodes in the taxonomy (twice - once with holes, once with Unclassified for the unknown/unreliable entries)
    ReduceOTUTaxonomy %taxoBLASTCutoffs% %OTU%_%type%_OTUTable.txt %OTU%_%type%_OTUTable_reduced.txt
    ReduceOTUTaxonomy %taxoBLASTCutoffs% -usX %OTU%_%type%_OTUTable.txt %OTU%_%type%_OTUTable_usX.txt

    rem Convert OTU to .biom format (name remapping string can be added as an extra parameter)
    ConvertTableToBIOM %OTU%_%type%_OTUTable.txt %OTU%_%type%_OTUTable.biom
    ConvertTableToBIOM %OTU%_%type%_OTUTable_usX.txt %OTU%_%type%_OTUTable_usX.biom

    rem summarise the OTU table at accession, species and taxonomy levels
    SummariseBLASTOTUTable %OTU%_%type%_OTUTable_reduced.txt %OTU%_%type%_OTUTable_summary.txt
    SummariseBLASTOTUTable %OTU%_%type%_OTUTable_usX.txt %OTU%_%type%_OTUTable_usX_summary.txt

    rem and generate results that only cover the OTUs that look like they belong to the target gene filtering
    if %filterOTUs% equ yes
        ThinOTUTable -fa %OTU%_%type%_only_OTUs.fa %OTU%_%type%_only_OTUTable_text %OTU%_%type%_only_OTUTable_summary
        ConvertTableToBIOM %OTU%_%type%_only_OTUTable.txt %OTU%_%type%_only_OTUTable.biom
        ReduceOTUTaxonomy %taxoBLASTCutoffs% %OTU%_%type%_only_OTUTable_text %OTU%_%type%_only_OTUTable_reduced.txt
        SummariseBLASTOTUTable %OTU%_%type%_only_OTUTable_reduced.txt %OTU%_%type%_only_OTUTable_summary.txt
        ReduceOTUTaxonomy -usX %taxoBLASTCutoffs% %OTU%_%type%_only_OTUTable_text %OTU%_%type%_only_OTUTable_usX.txt
        ConvertTableToBIOM %OTU%_%type%_only_OTUTable_usX.txt %OTU%_%type%_only_OTUTable_usX.biom
```
SummariseBLASTOTUTable %OTU%_%type%_only_OTUTable_usX.txt
%OTU%_%type%_only_OTUTable_usX_summary.txt
)

rem generate an OTU table that doesn't include the smallest OTUs
if %denoise% neq no (
DenoiseOTUTable %denoiseParams% %OTU%_%type%_OTUTable.txt
%OTU%_%type%_OTUTable_%denoise%.txt
ThinOTUs %OTUs %type%_OTUs.fa
%OTU%_%type%_OTUTable_%denoise%.fa
ConvertTableToBIOM %OTU%_%type%_OTUTable_%denoise%.txt
%OTU%_%type%_OTUTable_%denoise%.biom
ReduceOTUTaxonomy %taxoBLASTCutoffs%
%OTU%_%type%_OTUTable_%denoise%.txt
%OTU%_%type%_OTUTable_%denoise%_reduced.txt
SummariseBLASTOTUTable
%OTU%_%type%_OTUTable_%denoise%_reduced.txt
%OTU%_%type%_OTUTable_%denoise%_summary.txt
ReduceOTUTaxonomy -usX %taxoBLASTCutoffs%
%OTU%_%type%_OTUTable_%denoise%.txt
%OTU%_%type%_OTUTable_%denoise%_usX.txt
ConvertTableToBIOM %OTU%_%type%_OTUTable_%denoise%_usX.txt
%OTU%_%type%_OTUTable_%denoise%_usX.biom
SummariseBLASTOTUTable
%OTU%_%type%_OTUTable_%denoise%_usX.txt
%OTU%_%type%_OTUTable_%denoise%_usX_summary.txt
If %filterOTUs% equ yes (
DenoiseOTUTable %denoiseParams%
%OTU%_%type%_only_OTUTable.txt
%OTU%_%type%_only_OTUTable_%denoise%.txt
ThinOTUs %OTUs %type%_OTUs.fa
%OTU%_%type%_only_OTUTable_%denoise%.fa
ConvertTableToBIOM
%OTU%_%type%_only_OTUTable_%denoise%_reduced.txt
%OTU%_%type%_only_OTUTable_%denoise%_summary.txt
ReduceOTUTaxonomy %taxoBLASTCutoffs%
%OTU%_%type%_only_OTUTable_%denoise%_reduced.txt
%OTU%_%type%_only_OTUTable_%denoise%_usX.txt
ConvertTableToBIOM %OTU%_%type%_only_OTUTable_%denoise%_usX.txt
%OTU%_%type%_only_OTUTable_%denoise%_usX.biom
SummariseBLASTOTUTable
%OTU%_%type%_only_OTUTable_%denoise%_usX.txt
%OTU%_%type%_only_OTUTable_%denoise%_usX_summary.txt
)}

echo copy results files
copy %OTUs %type%_OTUs.fa %startingDir%
copy %OTUs %type%_OTUtable.txt %startingDir%
copy %OTUs %type%_OTUtable.biom %startingDir%
copy %OTUs %type%_OTUtable_reduced.txt %startingDir%
copy %OTUs %type%_OTUtable_usX.txt %startingDir%
copy %OTUs %type%_OTUtable_usX.biom %startingDir%
copy %OTUs %type%_OTUTable_summary.txt %startingDir%
copy %OTUs %type%_OTUtable_usX_summary.txt %startingDir%
copy %OTUs %type%_OTUs.aln %startingDir%
if %filterOTUs% equ yes {
    if not exist %startingDir\%type\_only ( {
        mkdir %startingDir\%type\_only 
    }
    copy %OTU\_%type\_only_OTUs.fa %startingDir\%type\_only
    copy %OTU\_%type\_only_OTUTable.txt
    %startingDir\%type\_only
    copy %OTU\_%type\_only_OTUTable_reduced.txt
    %startingDir\%type\_only
    copy %OTU\_%type\_only_OTUTable_usX.txt
    %startingDir\%type\_only
    copy %OTU\_%type\_only_OTUTable.biom
    %startingDir\%type\_only
    copy %OTU\_%type\_only_OTUTable_usX.biom
    %startingDir\%type\_only
    copy %OTU\_%type\_only_OTUTable_summary.txt
    %startingDir\%type\_only
    copy %OTU\_%type\_only_OTUTable_usX_summary.txt
    %startingDir\%type\_only
    if exist %OTU\_%type\_discards_OTUs.fa ( {
        copy %OTU\_%type\_discards_OTUs.fa
        %startingDir\%type\_only
        copy %OTU\_%type\_discards_OTUs.aln
        %startingDir\%type\_only
        copy %OTU\_%type\_discards_PhiX_OTUs.aln
        %startingDir\%type\_only
        copy %OTU\_%type\_discards_OTUs.uc
        %startingDir\%type\_only
        copy %OTU\_%type\_discards_PhiX_OTUs.uc
    }
    ) }
    if %denoise% neq no ( {
    if not exist %startingDir\%denoise\ ( {
        mkdir %startingDir\%denoise\ 
    }
    copy %OTU\_%type\_OTUs\_denoise\.fa
    %startingDir\%denoise\ 
    copy %OTU\_%type\_OTUTable\_denoise\.txt
    %startingDir\%denoise\ 
    copy %OTU\_%type\_OTUTable\_denoise\.biom
    %startingDir\%denoise\ 
    copy %OTU\_%type\_OTUTable\_denoise\_reduced.txt
    %startingDir\%denoise\ 
    copy %OTU\_%type\_OTUTable\_denoise\_usX.txt
    %startingDir\%denoise\ 
    copy %OTU\_%type\_OTUTable\_denoise\_usX.biom
    %startingDir\%denoise\ 
    copy %OTU\_%type\_OTUTable\_denoise\_summary.txt
    %startingDir\%denoise\ 
    copy %OTU\_%type\_OTUTable\_denoise\_usX_summary.txt
    %startingDir\%denoise\ 
    if %filterOTUs% equ yes ( {
        copy %OTU\_%type\_only_OTUs\_denoise\.fa
        %startingDir\%denoise\ 
        copy %OTU\_%type\_only_OTUTable\_denoise\.txt
        %startingDir\%denoise\ 
        copy %OTU\_%type\_only_OTUTable\_denoise\.biom
        %startingDir\%denoise\ 
        copy %OTU\_%type\_only_OTUTable\_denoise\_reduced.txt
        %startingDir\%denoise\ 
        copy %OTU\_%type\_only_OTUTable\_denoise\_usX.txt
        %startingDir\%denoise\ 
        copy %OTU\_%type\_only_OTUTable\_denoise\_usX.biom
        %startingDir\%denoise\ 
    }
    )
}
copy %OTU%_%type%_only_OTUTable_%denoise%_usX.biom
%startingDir%\%denoise%
copy
%OTU%_%type%_only_OTUTable_%denoise%_summary.txt
%startingDir%\%denoise%
copy
%OTU%_%type%_only_OTUTable_%denoise%_usX_summary.txt
%startingDir%\%denoise%
)
)
rem clean up temp files
echo delete temporary files
del *.fa
del *.txt
del *.biom
del *.uc
if exist *.aln {
   del *.aln
}

cd /d %startingDir%
time /t
Appendix II: A critical review of the bacterial baptism hypothesis and the impact of caesarean delivery on the infant microbiome

A Critical Review of the Bacterial Baptism Hypothesis and the Impact of Cesarean Delivery on the Infant Microbiome

Lisa F. Stimson*, Matthew S. Payne and Jeffrey A. Koelen
Division of Obstetrics and Gynaecology, Faculty of Health and Medical Sciences, The University of Western Australia, North, WA, Australia

Numerous studies suggest that infants delivered by caesarean section are at a greater risk of non-communicable diseases than their vaginal counterparts. In particular, epidemiological studies have linked cesarean delivery with increased rates of asthma, allergies, autoimmune disorders, and obesity. Mode of delivery has also been associated with differences in the infant microbiome. It has been suggested that these differences are attributable to the "bacterial baptism" of vaginal birth, which is bypassed in cesarean delivery, and that the abnormal establishment of the early-life microbiome is the mediator of later-life adverse outcomes observed in cesarean delivered infants. This has led to the increasingly popular practice of "vaginal seeding": the intravaginal transfer of vaginal microbiota to the neonate to promote establishment of a "normal" infant microbiome.

In this review, we summarize and critically appraise the current evidence for a causal association between cesarean delivery and neonatal dysbiosis. We suggest that, while cesarean delivery is certainly associated with alterations in the infant microbiome, the lack of exposure to vaginal microbiota is unlikely to be a major contributing factor. Instead, it is likely that induction for cesarean delivery, intrapartum antibiotic administration, absence of labor, differences in breastfeeding behaviors, maternal obesity, and gestational age are major drivers of the cesarean delivery microbial phenotype. We, therefore, call into question the rationale for "vaginal seeding" and support calls for the halting of this practice until robust evidence of need, efficacy, and safety is available.

Keywords: cesarean delivery, vaginal delivery, delivery mode, neonatal microbiome, infant microbiome, developmental origins, vaginal seeding

BACKGROUND

Cesarean section (CS) delivery can be a life-saving procedure. However, globally, rates of elective CS delivery are increasing. In Organization for Economic Co-operation and Development countries, more than a quarter of infants are now born by CS (1). The increasing popularity of elective CS delivery has coincided with a rising prevalence of non-communicable diseases, prompting investigations into a possible causal link between the two.

Abbreviations: CS, cesarean section delivered; DNA, desoxyribonucleic acid; GBS, group B Streptococcus; IAP, intrapartum antibiotic prophylaxis; MRSA, methicillin-resistant Staphylococcus aureus; NICU, neonatal intensive care unit; NICU, organization for economic co-operation and development; qPCR, polymerase chain reaction; qPCR, quantitative polymerase chain reaction; RNA, ribonucleic acid; RNA, ribosomal ribonucleic acid; tCO, real-time quantitative polymerase chain reaction; NICU, socio-economic status; VDB, vaginally delivered.